GrOW Formative Evaluation for Mid-Term Review

Final Report

EVALUATION REPORT | FEBRUARY 2017

UNIVERSALIA

Universalia Management Group
245 Victoria Avenue, Suite 200
Westmount, Montreal, Quebec
Canada H3Z 2M6

www.universalia.com

Executive Summary

Introduction to the GrOW Program

GrOW is a five-year, partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC). Launched in 2013 with a focus on low income countries in sub-Saharan Africa and South Asia, the program aims to support policies and interventions that improve women's economic empowerment and contribute to societal well-being. Thirty research institutions in over fifty countries have been supported to strengthen the body of evidence, improve policy uptake and build in-country research capacity.

Mid-term Evaluation

This Mid-term Evaluation is a formative evaluation to assess the program's positioning for building a body of quality evidence, policy uptake and built in-country researcher capacity. Assessment was conducted at both research project level (including sub-studies) and at the overall program level.

Findings

Overall GrOW program positioning to achieve main outcomes

The value-added of GrOW has been its diversity in academic-practitioner, North-South, mixed-methods research, including program commitment to balancing rigour, policy uptake and capacity building. Such diversity has also, arguably, been the challenge for coherence.

GrOW is moderately well-positioned to achieve it outcomes, more than might be evident from an examination of the parts. In spite of a range of quality in individual research projects, there is a strong body of evidence emerging. Key policy narratives are emerging related to inclusive growth that fill key gaps in the WEE sector though how these narratives are presented at the program level will be important. Key WEE stakeholders and the literature caution that structural barriers persist in spite of modest gains. How GrOW positions its macro-economic growth narrative against or alongside other non-economic policies and interventions will be critical, indeed, political and influential.

The body of evidence is strong around constraints and opportunities related to occupational and sectoral segregation, women's role in non-traditional sectors such as mining, the relationship between economic growth and gendered participation in the labour market, the effect of gender equality on growth, considerations and opportunities in the care economy. There is context-specific evidence on how to address particular barriers such as skills training in mobility-restricted contexts, child care. Depending on quality by the end of the program, there may be additional evidence related to pathways to work and decent work, the relationship between paid and unpaid work, effects of cash transfers, early marriage, early childbirth on WEE.

At this point in the program, emphasis on linking contextual evidence with broader trends (macro-micro links) will be critical to the external validity of program level policy narratives. A sectoral policy focus and reconsideration of deliverables, could position the overall portfolio as one of moderate to high quality, policy uptake and capacity building.

Capacity building is widely evident, both for in-country researchers and institutions, according to interviews and assessment though hardest of the outcomes to detect.

Research project positioning for quality, uptake, built in-country research capacity

Based on assessment of rigour, legitimacy and important, research projects fell broadly into three groupings: strong across outcomes; could be improved with corrective mechanisms; relatively weak across outcomes. Within these groupings, a couple of other elements help to explain performance: governance structure and strength of macro-micro links. For the governance structure, several aspects came into play including: geographical diversity; nature of institutions involved/diversity in terms of academic-practitioner mix; number of institutions involved.

Elements of highly rated research projects and strategies

Highly rated research projects had strong coherence between qualitative and quantitative methods, clarity on policy uptake and good governance marked by complementarity of partners, a mix of scientific rigour and context knowledge with at least one partner having a track-record for evidence-based policy research. These projects also demonstrated high and public productivity in instruments, outputs, websites. The governance structures that fell more into this category were North-South mixed partnerships and North-south academic consortiums.

Elements of research projects experiencing challenges

Research projects experiencing challenges lacked a coherent theory of change as well as alignment between qualitative and quantitative analysis or macro and micro analysis. Two governance structures seemed to fall into this category more: multi-country mixed North-south consortiums and South-based single institution or partnerships.

There are project and program levers, corrective measures, that can support most weaker research projects to come to an acceptable level of quality, policy uptake and capacity-building.

Findings related to key program levers

Research Project Management

Team and governance structure mattered to ensure complementary skills. Essential skills included solid mixed-methods research experience and evidence-based policy research and engagement. External supports and focused capacity building may be required for some projects combined with a strong process of standards, guidelines, frequent dialogue and iteration.

Program oversight and mechanisms

Standards, close accompaniment, especially for weaker projects, and outcome-focused documentation played a key role in ensuring quality research and policy uptake. Stronger mechanisms such as withholding funds were not used. An overall approach of dialogue and learning was also important. Accountability and a culture of learning can be effectively brought together going forward.

Key Recommendations

To GrOW Program Management and Executive Committee

1. Segment emerging policy narratives and stakeholders
2. Do less with a focus on quality. Reconsider deliverables to match type of research and differentiate outputs to match emerging policy narratives
3. In future calls for proposals, consider how design might support coherence for mixed-methods and macro-micro links in research

To GrOW Program

4. Continue to share what is working well
5. Focus on gaps in the WEE sector and program-level policy narratives going forward
6. Encourage projects having difficulties to focus more to align macro-micro links and ensure evidence feeds into program-level policy narratives
7. Balance differentiation in policy influence and capacity needs with accountability
8. Focus on outcomes for policy uptake and capacity building

Lessons Learned

About research quality and evidence

The GrOW program, in its diversity and complexity, is a microcosm of the broader field. As such, tensions and debates here are helpful to reflect on. What is the role of research in redressing power dynamics? The program demonstrates that the three outcomes can be mutually reinforcing but cannot be assumed. While the assessments confirm that a minimum of rigour is necessary to have policy uptake and capacity building, it cannot be assumed that rigour leads to the other two outcomes. As many, including in-country researchers emphasized, building capacity well takes time, resources, intentional research process design. It is critical, given the power dynamics and gender norms at play in this area, that the questions and researchers continue to be diverse in all of the backgrounds that intersect with gender.

About monitoring and evaluation

Outcome mapping within a systems approach to monitoring and evaluation supports not only an attention to results and assessments but to dialogue around perceived processes and assumptions related to change and influence.

Conclusions

The GrOW Program is positioned to add value and address gaps in the broader field of Women's Economic Empowerment related to inclusive growth. To do so requires strong and strategic focus on the overall GrOW Program and careful positioning of the message that will potentially have political implications and influence. It is critical that each project is focused on where they can best contribute to the larger body of evidence, policy narratives and dialogue.

Acronyms

CMO	Context, Mechanisms, Outcomes
DFID	Department for International Development
FEMNET	The African Women's Development and Communication Network
GrOW	Growth and Economic Opportunities for Women
IDRC	International Development Research Centre
IOA	Institutional and Organisational Assessment
MTE	Mid-term Evaluation
ODI	Overseas Development Institute
UN	United Nations
UN Exxon	United Nations Foundation - ExxonMobil
WIEGO	Women in Informal Employment: Globalizing and Organizing

Contents

Figures

Tables

Appendices

1 Introduction to the GrOW Program

1.1 Overview of GrOW Program

GrOW is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC). Launched in 2013 with a focus on low income countries in sub-Saharan Africa and South Asia, the program aims to support policies and interventions that improve women's livelihoods and contribute to societal wellbeing. The first component of the program approved eleven projects (of which three are multi-country) addressing the barriers to women's economic empowerment and the gender gaps in earnings and productivity. The second component of the program approved four projects (of which all are multi-country) on the effect of specific patterns of economic growth on women's economic empowerment. The third component approved one project (which is multi-country) on the effects of women's economic empowerment on economic growth.

1.2 Profile of Research Projects

Table 1.1 below provides an overview of the 14 projects that were reviewed for the mid-term evaluation.

Table 1.1 Profile of Research Projects

TITLE OF PROJECT	SITES OF STUDY	(LEAD) AND PARTNERSHIPS GOVERNANCE STRUCTURE	BUDGET ($CAD) AND TIMEFRAME	METHODS	EARLY OUTPUTS (Y/N)	REPORTING COMPLETE (Y/N)
107807 – Examining women's early labour market transitions in sub-Saharan Africa	Kenya, Burkina Faso, Ethiopia, Ghana, Tanzania, and Uganda.	(University of Sussex) African Economic Research Consortium (Mixed North-South Consortium)	777,600 34 months	Quantitative: census; household survey Qualitative: life history	Y 6 working papers	Y

TITLE OF PROJECT	SITES OF STUDY	(LEAD) AND PARTNERSHIPS GOVERNANCE STRUCTURE	BUDGET ($CAD) AND TIMEFRAME	METHODS	EARLY OUTPUTS (Y/N)	REPORTING COMPLETE (Y/N)
107808 – Improving childcare options to create better economic opportunities for women in Nairobi slums	Korogocho slum in Nairobi	(McGill University) Africa Population Health Research Centre (Mixed North-South Partnership)	952, 466 36 months	Quantitative: RCT Qualitative Qualitative/participatory: photo-voice	Y Briefing Note	Y
107809 – Understanding the effects of skills training on women's economic opportunities in Pakistan	Rural Punjab, Pakistan	(Harvard University – Kennedy School of Government) Centre for Economic Research in Pakistan (Mixed North-South Partnership)	904,600 24 months	Quantitative: RCT Qualitative: focus group discussions; individual interviews	N	N
107811 – Assessing the impact of cash transfer programs on women's economic empowerment in Tanzania	Tanzania	(Research on Poverty Alleviation) REPOA (South Single Institution)	816,300 36 months	Quantitative: WEAI Qualitative: focus group discussions; individual interviews	Y Pre-analysis plan 2 working papers-qualitative; quantitative	Y
107818 – Together we can: the role of women's action groups as agents of social and economic change in India	Bihar and Karnataka States, India	(Centre for Budget and Policy Studies) Institute for Financial Management and Research (South-Based Partnership)	902,300 36 months	Quantitative: RCT Qualitative: individual interviews; ethnographic research	N	N

TITLE OF PROJECT	SITES OF STUDY	(LEAD) AND PARTNERSHIPS GOVERNANCE STRUCTURE	BUDGET ($CAD) AND TIMEFRAME	METHODS	EARLY OUTPUTS (Y/N)	REPORTING COMPLETE (Y/N)
107819 – The influence of affordable daycare on women's empowerment in India	Rajasthan, India	(McGill University) Institute for Financial Management and Research, Seva Mandir (Mixed North-South Partnership)	1,000,171 36 months	Quantitative: RCT	Y (Baseline report; Guide to measuring WEE; working paper; protocol for systematic review)	Y
107820 – Uncovering women's experience in artisanal and small-scale mining in Central and East Africa	Rwanda, Uganda, and the Democratic Republic of the Congo.	(Carleton University) Partnership Africa Canada; Development Research and Social Policy Analysis Centre (Mixed North South Consortium)	996,087 36 months	Quantitative: survey Qualitative: focus group discussions; individual interviews e:	Y (Book chapter draft; working paper)	Y
107821 – Addressing the barriers to young women's economic empowerment in Bangladesh	Rural Bangladesh	(Innovations for Poverty Action) Duke University (Partnership)	970,700 36 months	Quantitative: panel data Qualitative: focus group discussions; individual interviews	Y (Draft qualitative journal article; cost-based analysis; working paper)	Y
107850 – Identifying post-war growth and economic opportunities for women in Sri Lanka	Northern Sri Lanka	(International Centre for Ethnic Studies) Point Pedro Institute of Development (Mixed North South Partnership)	719,000 36 months	Quantitative: household survey Qualitative: household case studies; focus group discussions; key informant interviews	Y (Literature review)	Y

TITLE OF PROJECT	SITES OF STUDY	(LEAD) AND PARTNERSHIPS GOVERNANCE STRUCTURE	BUDGET ($CAD) AND TIMEFRAME	METHODS	EARLY OUTPUTS (Y/N)	REPORTING COMPLETE (Y/N)
107852 – Balancing unpaid care work and paid work in South Asia and Sub-Saharan Africa	India, Nepal, Rwanda, and Tanzania	(Institute of Development Studies) Institute for Social Studies Trust; BRAC Uganda (Mixed North South Consortium)	992,700 27 months	Quantitative: household survey Qualitative: household case studies; focus group discussions; key informant interviews	N	Y
108111 – Pathways for shared prosperity: Understanding the link between women's economic empowerment and growth	Global, Pakistan, India, Jordan, Nepal, Colombia, South Africa	(Georg-August University Gottingen) Delhi School of Economics Stellenbosch University; University of Cape Town; Wageningen University (Academic North South Consortium)	2,255,300 24 months	Quantitative: Systematic reviews on experimental research; household survey; econometric/cross country Qualitative: focus group discussions; in depth interviews	Y (Systematic review protocol; 15 working papers)	Y
108112 – Making growth work for women in low income countries	Bangladesh, Brazil, Ghana, India, Kenya, Malawi, Morocco, Nigeria and Pakistan	Urban Institute (Mixed North South Consortium)	2,166,300 24 months	Quantitative: econometric/ cross-country Qualitative: systematic reviews	Y (Literature review; 2 working papers)	Y
108114 – Bringing to the light the role of the extractive industry on women's economic empowerment in Cote d'Ivoire and Ghana	Ghana and Cote d'Ivoire	(International Institute for the Advanced Studies of Cultures, Institutions and Economic Enterprise) Centre Ivoirien de Recherches Economiques et Sociales (South-based Partnership)	336,800 18 months	Quantitative: regression model/ econometric Qualitative: surveys/ interviews	N	N

TITLE OF PROJECT	SITES OF STUDY	(LEAD) AND PARTNERSHIPS GOVERNANCE STRUCTURE	BUDGET ($CAD) AND TIMEFRAME	METHODS	EARLY OUTPUTS (Y/N)	REPORTING COMPLETE (Y/N)
108115 – The impact of women's political participation on economic growth and women's economic empowerment in Africa	38 African countries	(University of Chicago) Stellenbosch University (Academic North South Partnership)	337,100 24 months	Quantitative: statistical analysis	N (no research outputs expected, only database)	N

2 Mid-term Evaluation

2.1 Universalia Management Group

Founded in Canada in 1980, Universalia (www.universalia.com) is a management consulting firm specializing in monitoring and evaluation, results-based management, performance measurement, organisational assessment, and capacity building. Universalia is a recognised leader in advancing the practice of evaluation through publications and the development of unique tools, such as the Institutional and Organisational Assessment (IOA) Model that was developed jointly with IDRC to explore the performance of research institutions and the factors that support or inhibit their performance.

Universalia offers services in a wide range of international development fields, including, inter alia : i) gender equality and empowerment; ii) micro-finance, livelihoods and economic development; iii) socio-economic policy, civil society and social inclusion; iv) education, youth programs and capacity building; v) health (including HIV/AIDS); vi) governance, security, justice, and human rights (including child protection); vii) water and sanitation; and viii) environment, agriculture and sustainable development. We have worked with a wide range of international, regional and national development organisations, including government agencies, UN agencies, research institutions, development banks, foundations, and NGOs such as IDRC, the MasterCard Foundation, the Rockefeller Foundation, UN Women, etc.

Headquartered in Montreal as a private incorporated company, the firm is owned by six shareholders and supported by a team of 20 permanent consultants and technical and accounting staff, providing services in English, French, Spanish, Arabic, Portuguese and German. Backed by Universalia's commitment to excellence, our team has over 55 combined years of experience in women's economic empowerment as well as solid research and research review expertise.

For this particular contract, Nanci Lee, the Mid-Term Evaluation Team Leader, brought over twenty years experience in Women's Economic Empowerment to the Mid-Term Evaluation. Her colleagues Rima Slaibi and Halcyon Louis brought combined experience of over twenty years in economic development, microfinance and financial literacy. Detailed Evaluation Team Biographies are found in Appendix II.

2.2 Mandate and Use of the Evaluation

The Universalia Management Group Limited was contracted by IDRC to conduct a formative mid-term evaluation of the GrOW project, of which fourteen projects remain supported.

This formative, Mid-term Evaluation examined all of the current portfolio of research projects funded through the Growth and Economic Opportunities for Women (GrOW) program by IDRC, DFID and the William and Flora Hewlett Foundation.

The conceptual and analytical framework was designed based on the research purpose of the Mid-Term Evaluation (MTE) of the GrOW Research Program. We understood the purpose as two-fold: **to assess the quality of research, effectiveness of project positioning strategies for uptake and built in-country researcher capacity; and the positioning of the overall program to achieve its outcomes by the end of the program.** Given that many of the fourteen research projects were still in process, the evaluation assessed early research outputs (or in their absence methodology and instruments) and the positioning of research projects to achieve policy uptake and built capacity. However, the main focus was an assessment of the overall program positioning to achieve its three key outcomes: strengthened body of evidence for WEE, enhanced policy uptake and built in-country research capacity.

The original Terms of Reference, the evaluation requested the evaluation to focus on (1) mechanisms in place for assuring research quality (2) strategies in place for enhancing research accessibility and for positioning research for uptake by scholarly, practice and policy communities; and (3) application of ethical research practice and integration of gender analysis.

While staying true to this initial emphasis, the framework for the evaluation focused on the three key outcomes of GrOW both at the research project level and the overall program level. Mainly the evaluation focused on positioning for: **quality** evidence; **uptake** of research by key decision-makers; built **capacity** of in-country researchers. Ethics was considered part of legitimacy which is an element or parameter within quality. Other elements of legitimacy included gender and contextual considerations as well as research team reputation. Accessibility was an element of the assessment of uptake.

GrOW had three separate thematic calls for research: 1) barriers to closing gender gaps and opportunities to address these barriers; 2) impact of patterns of economic growth on women's economic empowerment; and 3) contribution of women's empowerment to economic growth, all in varying stages of completion. Therefore, we focused on a mid-term formative evaluation that could provide concrete and actionable recommendations to research project officers to make improvements and adaptations to their projects and to the overall research program. While gender-responsiveness was specifically located within one of the dimensions, it is important to note that the whole evaluation process aimed to interrogate gender and gender components of the work.

The evaluation team supported the statement from the DFID Annual Review 2016 that:

> "IDRC plans to carry out a mid-term review which will include an assessment of the quality of research approaches. We suggest that the review is timed so that it can include assessment of a number of research outputs, including the working papers and policy briefs, and not just methodologies." (DFID, 2016, p.2)

The mid-term evaluation was more effective in supporting mid-term learning and adaptation through the assessment, at least in draft form, of some of the research outputs. Where this was not possible, methodology and instruments were requested. This was the case for four of the fourteen research projects.

2.3 Evaluation Questions

As per the Original Terms of Reference, the key evaluation questions were the following:

1) Effectiveness of GrOW Implementation: How effective and appropriate has IDRC's implementation of GrOW been, relative to the purpose and objectives of the program?
   - Did GrOW follow the value for money guidelines with regard to research specification and procurement as set out in the program design?[1]
   - What role if any has IDRC's approach[2] played in contributing to the program's effectiveness particularly with respect to the quality and relevance of research and research uptake? How effective has this approach been for GrOW?
   - What are the key lessons for GrOW to improve its effectiveness for the tasks to be undertaken in 2017-2018?

2) Research Quality: How effective have GrOW and GrOW grantees been in assuring research quality at this stage of the research process?
   - What are some examples of good practice?
   - Does a systematic assessment of methodology (both quantitative and qualitative) across GrOW projects provide confidence about the credibility of future outputs?
   - If there are any constraints on the methodological capacity of GrOW projects, how effectively have they been addressed?
   - How can GrOW and GrOW grantees strengthen research quality further?

3) Planning for Research uptake: To what extent have GrOW and GrOW projects put in place strategies and practices to position research for use by target audiences, such as influencing government policy or modifying practice by implementing agencies?
   - What are some examples of good practice?
   - How effective has GrOW and GrOW project been in undertaking research uptake at the national and international level?
   - How could positioning for use be enhanced?

4) Ethical practice: How successfully are GrOW grantees applying acceptable research ethics and security practices in the implementation of their research projects?
   - What are some examples of best practice? What should be avoided?
   - How can ethical research practice in GrOW be strengthened?

5) Research products: Based on the review of research processes, methodology and the available interim research and research uptake products, what would the quality dimension and sub-dimension rubrics look like for GrOW?
   - How does the four inter-related dimensions of research quality: integrity, legitimacy, importance and positioning for use, inform an assessment of GrOW's expected performance?
   - How can the quality of interim and future GrOW research outputs be aggregated and assessed?

2.4 Methodological Approach

The methodology for this mid-term evaluation drew heavily on the conceptual framework of women's economic empowerment, which is understood through the pioneering gender work of Maxine Molyneux and Naila Kabeer. Empowerment, as Kabeer (2015) explains, is "the process by which those who have been denied the ability to make strategic life choices acquire such an ability. Women's economic empowerment, subsequently, is about much more than economics, labour markets and finance. Molyneux helped distinguish and make the links between practical (access to resources and opportunities) and strategic (voice in decision-making; control over resources; ability to take advantage of opportunities) gendered interests (Molyneux, 1985). AWID (2010) provides a framework that helps to show how these domains of change inter-relate and reinforce one another: societal gender norms and practices; interventions and access to, control over resources; policies, laws and budgets that provide rights and entitlements to resources and assets. Similarly, Kabeer (2015) provides a framework that reminds us to distinguish access to opportunities from outcomes and agency. In other words, the extent to which women really can take advantage of opportunities and exercise choice (Kabeer, 2015).

There are important links between informal economic activities (where women disproportionately reside) and the formal economy (Meagher/WIEGO, 2013) as well as local responses in a globalized economy. A gendered understanding of economic empowerment moves beyond the commonly understood areas of income, employment, markets, economic security to include access to land and property rights, unpaid, decent and informal work as well as social protection including cash transfers (International labour organization, 2012; Meagher/WIEGO, 2013; Kabeer, 2015). The evaluation approach used a conceptual framework for women's economic empowerment that was robust enough to speak to the wide-ranging realities and contexts in which the research is operating.

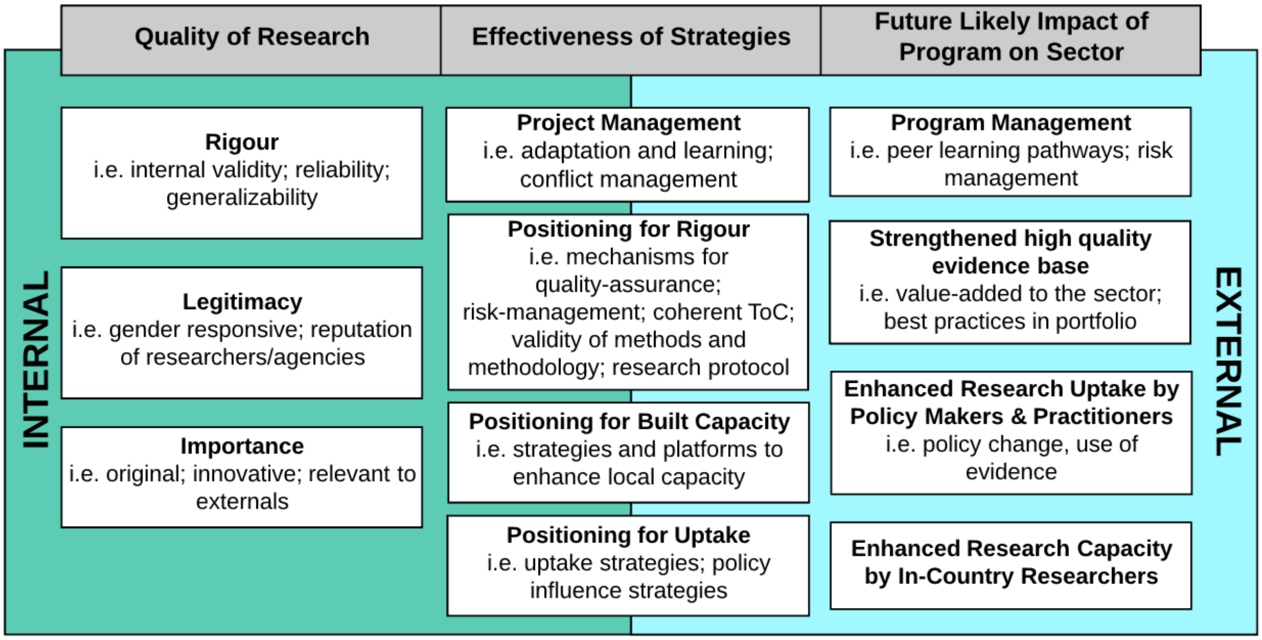
The methodology used a systems and theory of change approach focused on outcome mapping, pathways to change and the stimulation of dialogue between key stakeholders. This approach differs slightly from many approaches to evaluation that are log-frame based in that the emphasis is on changing perceptions and the dialogue as much as capturing progress and results. The approach also acknowledges that there are many pathways to change. Part of the role of a formative evaluation is to uncover tensions and trade-offs in how different stakeholders are positioned along these pathways.

In terms of parameters, this approach drew heavily on the IDRC Research Quality Plus paper (Ofir et al, 2016) for the domains as well as DFID's Quality of Evidence paper (2014) in terms of principles of high quality research and how to assess a body of evidence. As the diagram below shows, we organized the evaluation into three broad domains:

- Individual research project quality (in terms of positioning for rigour; legitimacy; relevance);
- Effectiveness of project positioning (in terms of project management, positioning for uptake and in-country research capacity); and
- Overall, the positioning of the research program (in terms of program management and to achieve the three main outcomes on quality evidence, uptake and built capacity).

The evaluation combined examination of the early research outputs (at varying stages of completion) with project strategies (management, positioning etc.) and perspectives of sector stakeholders.

Figure 2.1 Methodological Framework for Mid-Term Evaluation



As the diagram shows, if we move from left to the right of the framework we move from more research project control (internal) to more focus on the overall research program as well as influence on the sector as a whole as well as country-specific sectors (external). Key to this methodological approach is the understanding that quality can be assessed from three vantage points. On one hand, we could examine the research instruments, design and draft outputs independently. However, that only provides part of the picture. We could also assess the strategies and mechanisms that have been put in place to position the research project for ensuring rigour, uptake and in-country capacity. The most important aspects of these assessments related to capturing the trajectory and progress of the research projects regardless of their starting point or current level of completion. The aim was to assess the projects consistently not comparatively. The evaluation team aimed to assess using outcome mapping and narrative descriptions about the trajectory of the research projects and how they can be improved going forward.

Finally, the evaluation examined how well the research sits in the sector as a program overall. That is, given the formative nature of the mid-term evaluation, what is the current trajectory and likely impact of the overall research program? This could be determined by an assessment of the research program and its strategies as well as consultation with both internal and external stakeholders. An evidence mapping framework allowed the research program to be positioned within the overall sector of women’s economic empowerment.

2.5 Methods and Limitations

2.5.1 Research Project Quality and Effectiveness

Each research project was assessed in terms of both "stand-alone" quality of outputs as well as positioning for policy uptake and built capacity. In other words, an assessment based on the research itself irrespective of its positioning for policy uptake or built capacity which is assessed separately. Even though IDRC combines policy uptake and built capacity as part of its treatment of quality, the evaluation team chose to separate them for two reasons. First, positioning for policy uptake involves a long process of strategies that cannot be necessarily assessed from early outputs alone. Much more so than rigour, policy uptake and capacity building are part of a pathway so progress has to be assessed in terms of how well positioned the project is to achieve its outcomes by the end. Second, there are differences of opinion within the GrOW program and in the development community related to how to define quality research. In order to render these tensions transparent and analyze them well, it was necessary to unbundle the components for the evaluation.

With approval and iteration with IDRC, the original questions in the Terms of Reference were refined to the following key questions that guided the research project assessments in the Mid-Term Evaluation:

Research Quality of each Research Project

- **Rigour**. What is the rigour of the research project based on methodology, instruments and early research outputs? **Data sources:** GROW Document review; Review of methodology, instruments, draft research outputs
- **Legitimacy**. What is the legitimacy of the research in terms of its reputation, grounding in local context and attention to gender and ethics considerations? **Data sources:** Interviews with key sectoral stakeholders; Survey and interviews with principal investigators and lead researchers
- **Importance**. To what extent does the research add-value and fill a gap in the sector with respect to innovation or importance? **Data Sources:** Draft research review; Interviews with key sectoral stakeholders; systematic and literature reviews in women's economic empowerment
- **Synthesis**. What is the indication of quality by this research project, and its sub-studies, based on its current trajectory?

It is important to note that where the research project had many sub-studies, the analysis of quality was done at a composite level even if quality ranged across the sub-studies. In the detailed assessments, the evaluation team considered the variance between studies as well.

Effectiveness of Research Project Positioning

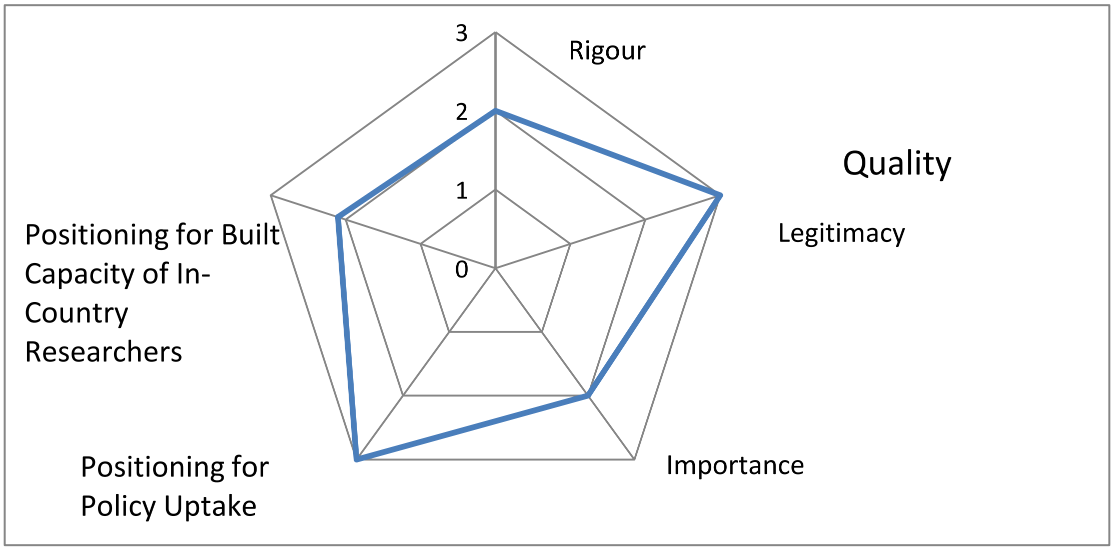
- **Project Management**. How effectively is the project managing to ensure quality, risk management and adaptation to emergent learning? **Data sources:** Document review; Interviews with project officers; Survey and interviews with principal investigators and lead researchers
- **Positioning for Uptake**. How effectively is the project positioning the research for uptake by key influential decision-makers? **Data sources:** Document review; Interviews with project officers; Survey

and interviews with principal investigators and lead researchers; Interviews with key sectoral stakeholders

- **Positioning for Building In-**Country Research Capacity. How effectively is the project creating opportunities for building in-country capacity? **Data sources:** Document review; interviews with project officers; Surveys and interviews with principal investigators and lead researchers
- **Synthesis**. How well positioned is the research project to achieve quality evidence, research uptake and built capacity of in-country researchers by the end of the program given its current trajectory?

Each of the domains and parameters (such as rigour) have their own outcome mapping statements (expect to see; like to see; love to see). Based on these, each research project as well as the overall program is plotted on a diamond plot. The purpose of the diamond mapping is for the program management and project managers to visualize some of the trade-offs and tensions that exist across domains and parameters without identifying projects or placing a judgment on those trade-offs.

Figure 2.2 Example of a Diamond Map Assessment of a Research Project



2.5.2 Overall Research Program Positioning

Synthesis of the overall Research Program was more than the combination of individual research project assessments. In fact, the assessment of the overall program was the most important part of the evaluation. The Overall Research Program was assessed based on its positioning to achieve the three key outcomes (DFID, Annual Review, 2015) and its program effectiveness toward:

i. Strengthened evidence base that is high quality and generated by top researchers from the Global North and South bringing together different expertise to answer enduring research questions

ii. Enhanced research uptake by policy-makers and other key influencers (academe, practitioners)

iii. Enhanced research capacity of in-country researchers in the design, conduct, management, dissemination and communication of rigorous, policy-relevant research.

Therefore, the overall body of evidence was assessed in terms of its quality, strength, value-added to the WEE sector particularly for policy influence as well as if the program has made a significant contribution to building in-country researcher capacity.

The Overall Research Program was assessed through the following data sources:

- Key sectoral representatives in women's economic empowerment who can triangulate analysis of research quality and speak to relevance and value-added (depending on scope, one for each region is desirable). For a draft list of potential stakeholders, see Tool 10;
- Review of key Research Program materials and documents such as DFID Program Reviews and Technical Reports including feedback and iteration; and
- Key systematic reviews and literature reviews on women's economic empowerment (within and outside of GROW) including reviews by UN Women, the ILO, Women and Development Network/AWID, ODI and IDRC's own literature review that began the program.

For more detail on the evaluation questions, indicators and data sources see Detailed Evaluation Matrix (Appendix II) and the various evaluation tools that were used during data collection, Tools 1 to 10 (Appendix V).

2.6 Phasing

The phasing of the Evaluation was done in three overlapping phases:

i. The Inception Phase

ii. The Individual Research Project Phase (including Mid-Term Workshop)

iii. The Overall Research Program Phase

Data collection and analysis took place largely in the second and third phase (though there was limited data collection in the first phase) with synthesis and final report prepared in the final phase.

I. Inception Phase

The main purpose of this phase was to clarify the broad scope, evaluation methodology and framework with IDRC/GROW (kick-off meeting) in order to develop the inception report (evaluation plan and detailed tools). The most pressing tool to finalize by the end of September was the survey for the Individual research projects to allow adequate time for receiving, analyzing and preparing for the Mid-Term Workshop in Germany. During this phase, the team also began to review the research and background documents to understand, and address early, the completeness of data and revision of evaluation scope. Initial interviews with project officers also began during this phase. This phase included a preliminary review of available program and project documentation to support contextual understanding of GrOW, including the general profile of individual projects.

II. The Individual Research Project Phase

During this second phase the majority of data collection and analysis was done to assess the quality and effectiveness of the individual research projects. In order to generate evaluation findings that provide an overview of project implementation activity, document review was conducted on all 14 projects with all available research outputs. For the six projects that are being implemented in multi-countries, in-country researchers were selected for interviews based on:

- Availability of draft working papers, policy briefs or preliminary results;
- Ensuring French-language speaking in-country researchers; and
- Ensuring a diversity of geographical locations

Meetings at the Mid-Term Workshop were key. The aim was to have completed the document review of several draft research outputs, as well as to have received the bulk of responses from the surveys prior to the Mid-Term Workshop. The Mid-Term Workshop was used to: to fill data gaps; to clarify, probe and explore issues in more depth that arose from other results and review. It was expected that some of these interviews and iteration with project managers would continue into the third phase. As well, additional research team members not present in Germany were identified for follow-up interviews.

III. The Overall Research Program Phase

The third and final phase was focused on any outstanding analysis and data gathering to understand the quality and effectiveness of the individual research projects. Here it was important to understand: i) how the different theoretical frames help to contribute to an impact narrative and body of evidence that adds value to the sector; ii) the interaction and tensions between quality of evidence, uptake and built capacity; and ii) the role of the IDRC and the research program overall to spur quality research, uptake and built capacity. The focus of this phase was therefore on understanding the Overall Research Program as more than the sum of its parts. During this phase, follow up interviews with Principal Investigators were done as well as interviews with selected in-country researchers. Synthesis of the previous analysis was completed. A strategic assessment of the effectiveness of positioning of the overall research program, including its emerging body of evidence, was done based on data available to date with a focus on formative recommendations and positioning for the future. The draft report will be completed during this phase with a Skype call following the submission. Dialogue and feedback provided, based on an assumption of no more than two rounds of feedback, will be used to complete the final report.

3 Assessment Results

3.1 Overall, Grow Program Moderately Well- Positioned to achieve main outcomes

The ultimate aim of GrOW is to contribute to increased uptake of policies, programs and interventions that improve economic outcomes and opportunities for poor women in low income countries. It seeks to do so through a strengthened evidence base, enhanced research capacity, and effective research uptake for policy. As assessment of both research projects and the overall GrOW Program have shown, GrOW is moderately well-positioned to achieve it outcomes. There is a strong body of evidence emerging that is linked to a number of policy narratives that fill key gaps in the WEE sector as identified by key WEE literature reviews and external stakeholders. Capacity building is evident, both for in-country researchers and institutions, according to interviews and assessment though least clear of the outcomes.

As intended, the GrOW Program will provide needed evidence on the relationship between economic growth and women's economic empowerment in both directions. In addition, there will be evidence to support a better understanding of factors that affect labour market entry, transitions to work and decent work, occupational and sectoral segregation and the relationship between paid and unpaid work. The main area that needs to be strengthened for the remainder of the program is the link between these broader trends and context-specific evidence (often at the household level) that helps to provide gendered contextualization and nuanced explanations to complement the macro-level data.

Assessment at the program level focused on a number of high-level literature reviews on Women's Economic Empowerment as well as interviews with number of key sectoral stakeholders representing a range of practitioners, academe, donors and multi-lateral technical experts, Naila Kabeer (London School of Economics), Caren Grown (World Bank), Abigail Hunt (ODI), Jennefer Sebstad (formerly USAID), Linda Jones (MEDA), Elizabeth Vasquez (WEE Connect) and Karen Stefiscyn (PowerAfrica Project, formerly Centre for Human Rights, University of Pretoria), Krista Jacobs (USAID), Srilatha Batliwala (CREA) and Sophia Mwakagenda (Member of Parliament, Tanzania, former gender activist). The evaluation team experience, in combination with interviews held with Principal Investigators and In-country researchers, also contributed to the analysis.

3.1.1 Key sectoral stakeholders emphasized policy narrative coherence and links between micro and macro-analyses

Overall, key sectoral stakeholders stress the legitimacy of the donor organizations and many of the research institutions, the networks and the importance of coherence for the policy agenda, particularly the links between macro and micro issues. They also offered some insights into the gaps they see and suggestions going forward.

Key sectoral feedback can be summarized in a few key themes that emerged:

- Key value-added is the alignment between GrOW program and high-level strategic foci of donors and their networks;
- The importance of a focused, coherent, program narrative for policy influence;
- The importance of linking micro-analysis to macro-analysis; and
- The desire to be well-informed about the research and engaged in dialogue.

Alignment of GrOW program with High-level strategic policy foci of donors key value-added

Stakeholders agreed that the donors bring high-level strategic foci including the policy networks in their fora. They emphasized the importance of the policy uptake focus and drawing on networks as is certainly being done. Stakeholders and PIs stressed that a key-value added of the GrOW program was the balancing of outcomes particularly policy uptake and capacity building. These additional aspects are often expected but not emphasized.

Many also commented on the importance of mixed method research and appreciated the structure of the North-South partnerships and consortiums and the richness this diversity presents both for evidence and policy influence. Many emphasized the importance of the GrOW approach to tackling such complex issues in WEE. The following are some examples of comments:

- *Solid research institutions and universities involved it looks like, good diversity. The North-South partnerships are important.*
- *The openness of the individuals working in the program. I really appreciate that and pushes the agenda and helps us to be better at the work we are doing. Both politically and policy wise. Research and evidence side.*
- *I think that model of having everyone in the room with some evidence and debating it is great.*
- *In the process, there are more policy oriented people involved. Definitely it seems to me that that is a good mix.*
- *With the limited information I have, I like the structure of the research. I think it is a really positive approach. The North-South, and mixed methods. I think qualitative research is very strong in helping us get the nuance.*
- *As far as I know IDRC puts a great deal of emphasis of doing policy engagement from the start. You get interest with the kinds of people who might be interested. That is not typical academic practice.*
- *Rarely do you have the luxury of building policy influence and capacity building into the design*

Importance of a coherent program strategy

One of the most common statements by external stakeholders was the importance of coherence for policy uptake and impact of the program overall. While many admitted to not knowing the research issues in detail, based on the summary of the projects they found the range of projects to be broad. As has been discussed internally in the program, the diversity was one of the intentions and has been both positive and with challenges. The impression of one stakeholder is described below:

- *Different outcomes and different interventions. Not necessarily a bad thing unless it is spreading itself so thin, with small samples across a diversity of markets, countries so it makes it difficult to do* *comparative analysis and trending. To the extent the program could increase investments in particular areas better and more data collection that can be compared across markets and time. Otherwise, you'll end up with very little bits of data and may or not be teachable moments for other practitioners very specific to those contexts.*

Other stakeholders had similar observations:

- *Mixed bag. Some seems extensions of existing work. Not sure how new or innovative. What is the new piece?*
- *Not building on one another. What, is the program wanting to influence? Not focused enough to really be able to contribute to a particular area? This is just too disperse. Seems to be quite a mix.*

On where the value-added was, there was a mixed response. While some described the connections between WEE and economic growth as the main value-added, others felt that there is a lot of macro-trend data and emphasis on experimental research. The following are examples of comments:

- *We have to move beyond relying on regressions that take no account of variations across country. They have to rely on the data that already exists. Can't explore new hypothesis and data.*
- *I have a real objection. The majority is experimental research. And they do not engage in mixed methods. They want to go in and do their experiments. Do they get researchers who know about that country?*
- *The research fills an important gap. There is a lot of debate about the relationship between WEE and economic growth. It will be good to have some solid evidence.*
- *I was formerly more of a qualitative researcher. But I've been more exposed lately. Now I see the value of quantitative research, intervention research. We need the mix. Mixed methods really allows us to get the full picture. Helpful to position the work around the relationship with economic growth. Timely.*

Importance of Linking Micro to Macro

There was a lot of consensus in the comments that it is important to make strong links between the micro-level and macro-level studies. Macro-studies are reliant on sub-optimal data, as one stakeholder commented. They need the contextual grounding of micro-level studies. Micro-studies, in contrast, can be too isolated or anecdotal. It is helpful to understand where they fit into larger issues and global trends. Also, they emphasized the importance of the program being clear what overall findings they will be able to glean in terms of these macro-micro connections. Some offered questions and others some suggestions based on their experience:

- *What is macro vs what is micro? Macro- GDP, sectoral issue. It could be a firm specific issue. There is a lot of micro stuff that is not so relevant to the macro conversation. So how does this GrOW competition distinguish itself to be macro?*
- *…Link projects to what is happening at the population level based trends.*
- *We had a review of similar types of research. The conclusion was that the research did not really get at the nuance. They couldn't explain the interesting whys and nuances. Why it showed up in one country and not in another. Also, looking at the cost of the interventions relative to the outcomes. That is coming up in discussions. Great if it is- how much it is costing when you are on the ground doing research.*

- *A tighter portfolio might be able to answer the gaps better.*
- *Policy briefs for that specific context. But will they pull out the key overall findings at the program level?*

Gaps

Stakeholders felt that there is a good mix of diversity. Some felt there may be too much diversity to yield coherent policy messages and evidence. Some also offered some observations in terms of gaps in the WEE sector overall that they are seeing in their work. Of course, no program could be expected to fill all of the gaps in such a complex and diverse sector. However, understanding their perspectives on gaps may help in terms of some of the project and program emphases in findings and messaging. The main gaps they identified were: non-poor women and policies that affect poor women; macro public policy; MENA region and global mechanisms. The following are some examples of comments:

- *We always seem to focus on the bottom of the pyramid. We need women in the middle hiring too. We have to. They are in the best position to hire women for good quality jobs. All important. Too often we focus only on the start-ups or the base. We need to do a lot of work along the value chain to make it equal opportunity.*
- *Land rights is a key issue that too often gets left out of discussions of economic opportunities and growth.*
- *They didn't fund the macro –public finance piece. They have funded open economy macro. Nothing on macro finance- exchange rates, currency issues. Inflation monetary. Bank questions.*
- *[to the question of gaps]….comparative work with different economic structures…. I was impressed by the geographic spread. Latin America left out a lot partly because they have their own body of research and done in Spanish. More highly focused on Sub-Saharan Africa, South Asia. MENA not there. Eliminates many historically Muslim countries.* [While it was explained that Sub-Saharan Africa and South Asia were the geographical focus, more than one stakeholder felt that MENA is important, politically, to include and has lessons to share.
- *Something looking at social norms around gender roles. What do we know? We know more about access than we do about agency.*
- *The huge link with sexual and reproductive rights. Engaging men. Expectations about women's reproductive role. Earning differentials and links to violence. It is trendy to talk about engaging men but there is not a lot of evidence.*
- *To what extent in a rights approach? Policies and laws. Are we stopping at the national level or are we looking at advancing those conventions at the global? CEDAW. The international framework obliges by international law, nations to do so. Child care, for example.*

One stakeholder asked some very helpful questions related to positioning the program policy narratives.

- *What is the framework that underlies the program? What is the understanding of what constitutes WEE that the program is building toward? I worry about a narrow interpretation about the relationship between WEE and economic growth. What are regional trends? I want a meta framework that can travel across these contexts sufficiently robust to be portable. Where is it bringing in global market forces? Huge demand for women's labour as migrant workers- nannies, domestic workers, setting aside trafficking for a moment which is also part of it. Both empowering and not. I would like to know. Somewhere the study has to take cognizance of the fact that evidence-based policy making* *with local policy makers. I'm sorry but that is a bit of a joke. You have to reframe and articulate WEE in this larger reality.*

Suggestions

The main suggestion was to network and communicate the findings well. Many stakeholders commented on how important it is to have clarity on the overall program aims and findings as they come out. They all expressed an interest to be on an email list-serve or some form of update, as a means of sharing the program findings:

- *Didn't realize that the research was not only focused on the relationship between economic growth and women's economic empowerment, also barriers, constraints and opportunities. Better publicize that if that is the case. GROW is misleading. If they want to bring others on the board. May be only those in economic development drawn to the program. If they are generally trying to have the breadth. So as not to lose interesting audiences.*
- *The thing that strikes me when I saw this program is that I have never heard of it. So relevant to AID and other donor agencies.*
- *Is there [a] newsletter? There should be. It is also a way of seeing the sum total. Not the individual projects but the overall story. What is the narrative of the big picture?*

3.1.2 Literature Reviews on Gaps in Women's Economic Empowerment Sector consistent and broader than GrOW focus

The GrOW Program has significant potential to contribute to the body of evidence in the Women's Economic Empowerment Sector if it can help to confirm, nuance and contextualize the evidence that exists between economic growth and women's economic empowerment. Not surprisingly, the gaps identified are broader than the GrOW Program. Nevertheless, the consensus around gaps in the sector is important to note as it will help to situation the next section that focuses on GrOW's emerging body of evidence on inclusive growth and how it can be strengthened.

This section summarizes literature and evidence reviews on women's economic empowerment which is broader than but consistent with GrOW's focus including GrOW's own literature review on and the concept paper commissioned to Naila Kabeer on WEE and inclusive growth (Kabeer, 2012; GrOW, 2013). There are literature reviews on WEE included by the Overseas Development Institute (2016), The UN Women report on Progress of the World's Women: Transforming Economies, Realizing Rights (2016); The International Labor Organisation, Women in Work (2016), UN Foundation and Exxon Mobil, WEE Pathways (2016); the Gender and Development Network/Association of Women's Rights in Development. There was quite a lot of consensus in their identified gaps and priorities:

- Transforming work to ensure it is decent and balanced with unpaid care work;
- Context-specific evidence on what works where, for whom and why;
- Structural issues, harmonized macro-economic and social policies; and
- Accountability that goes beyond borders.

Transforming work

The literature reviews were consistent in their identification of priorities for policies and interventions:

- **Tackling the root causes of sectoral and occupational segregation** (addressing discrimination; demand-driven skills training including non-stereotypical fields for both women and men; promotion of women's entrepreneurship, participation and leadership in decision-making)
- **Addressing the gender wage gap** (through limiting long paid hours and overwork, changing attitudes towards unpaid care work, preventing and eliminating discrimination, promoting equal pay for equal value through wage transparency, gender-neutral job evaluation methods)
- **Implementing a comprehensive framework to achieve harmonization of work and family responsibilities** (making quality and early childhood care and education a universal right; creating and protecting quality jobs in the care economy; ensure the provision of basic infrastructure especially in rural areas; policies and laws related to adequate social protection, maternity protection, gender-based leave policies; progressive taxation; family-friendly working arrangements)
- **Women's organizing and participation in decision-making** *(*from international agreements to intra-household decision making) economic decision making is political process including collective action and women's movements as well as formal representation.

(GrOW, 2013; ILO, 2016 a, p. 94-95; ILO, 2016b; Gender and Development Network, 2016; UN Women, 2016; Buvinic et al, 2016; Kabeer, 2012).

While the ILO and UN Women stressed the first three priorities with decision-making as a sub-component, the Gender and Development Network isolated both organizing and decision-making as its own priority.

> Women's organizing and the strength of their autonomous movements are the strongest predictors of gender equality laws and policies across a range of areas from family law to violence against women and from non-discrimination in employment to childcare services (UN Women, 2016, p.17).

Both the UN reviews and that of The Gender and Development Network also place more emphasis on asset building, not only income and employment. Asset building and entitlements are an important part of building resilience for sustained work and income. Formal ownership and control over land was identified as key issues particularly for rural and agricultural areas but also as important form of economic security elsewhere. Success of land tenure interventions depend on paying attention to social and local context. (UN Foundation, Exxon, 2016; Gender and Development Network, 2016; UN Women, 2016).

The GrOW Program is perhaps best positioned to contribute to these areas though land and asset entitlements may be a notable gap. The Sri Lankan work has potential to fill this gap if corrective measures can support a focus on what conflict and post-conflict has done to gendered asset entitlements and agency including land.

Context specific evidence

Key evidence reviews shared the same perspectives on evidence gaps in WEE, also on what is well known. Though more cross-country evidence is helpful, barriers and constraints to WEE are quite well-known and documented. There is a large and fairly consistent body of evidence that show that gender gaps exist worldwide in terms of employment, income, and especially decent work in spite of modest gains in access to education (UN Women, 2016, Buvinic et al., 2016; ODI, 2016). More evidence is needed to better understand what works, where and why in spite of or against these trends.

Context specific evidence helps to make links between micro level interventions and macro level policy, market and socio-cultural contexts. It helps to draw relationships between conditions for success. For instance, the UN Women report found “among developing regions, Latin America has seen the most progress in family-friendly policies over the past decades and has also seen the most significant increase in women’s labour force participation” (UN Women, 2016, p. 14). ODI (2016) in a review of methodologies for WEE assessed all of the evaluations that they reviewed, in terms of whether there was a sound gendered contextual analysis on which to build their design.

Using the region of Africa as an example, a comprehensive report by Civic Society based on 51 reports, including not only interventions and programs but policies, economies, legal frameworks, commitment to Millennium Development Goals and compliance with international conventions places context-based evidence high on their priorities (FemNet, 2016). They found four main contextual factors that affect implementation worth noting that act together: high-level ministerial or Presidential leadership; political stability; economy citing those with higher levels of growth and greater economic diversity as better placed to address gender equality; and finally, history. With respect to history, they noted that countries that have historically strong women’s movements as part of independence liberation struggles where women have been key to the nation making project and are recognized in public narratives, there tends to be greater political accountability and gender equality. The report also reported a lack of disaggregated data across the region to understand what existing laws and policies in place have yielded. They argue that evidence on progress without disaggregated data then remains anecdotal and open to influence by non-state actors (FemNeT, 2016, p16).

Reinforcing views of the external stakeholders, the literature emphasized how difficult but important it is to transform persistent social norms and to understand how this is possible. The UN Women report (UN Women, 2016, p. 24) prioritizes action on three interrelated fronts: redressing socio-economic disadvantage; addressing stereotyping, stigma and violence; and strengthening agency, voice and participation (UN Women, p24). The reviews shared the need to understand context-specific and gendered social norms that maintain these inequities even where access is improved.

There was also demand for evidence that showed how interventions and policy levers need to be adapted to accommodate different populations.

> Income and control of assets continue to lag behind those of men. This inequality between women and men also intersects with the substantial gap between rich and poor, as well as other discriminations women face on the grounds of race, age, disability, sexuality and so on. Recognising the intersectionality of barriers facing different women is essential for effective policymaking and achieving equality, as is an acknowledgement of the footprints of colonial history in shaping economic policies that exacerbate gender inequality (Gender and Development Network, 2016, p. 3).

The UN Foundation/Exxon Mobil study (2016) provided different road maps or pathways for WEE given an analysis of the mix of policies, interventions and different kinds of economic structures and population segments. They segmented the populations into: poor entrepreneurs; non-poor entrepreneurs; wage workers; poor farmers; non-poor farmers and young and highlighted promising, proven, high potential and unproven interventions. All of the reviews cited the critical importance of access to and control over assets.

"Integration into the markets require secure access to land and other resources. If not addressed up front, people lose claims as other opportunities increase. Depending on the context, different approaches need to secure land rights (UN Exxon, p. 96)." The roadmap was based on a typology of economic structures distinguishing highly fertile agrarian economies, declining fertility urbanizing economies, missing adults/conflict-affected economies and resources rich economies and small nations.

While these typologies and segments can be debated, they raise the important issue of how we understand and differentiate outcomes and interventions. What can be learned from these differences? What differences matter? What, if anything, needs to accompany economic growth to ensure that it is inclusive?

What will be learned in GrOW about different economic structures, different types of interventions, different pathways across contexts? The context-specific evidence is perhaps where the GrOW portfolio can most be strengthened. Without having strong dialogue and iteration with the research projects at this stage in the program cycle, there is a cluster of macro-studies related to macro-economic growth with quantitative trend data and another of micro-studies that are more country-specific or even more specific to household dynamics. As GrOW's earlier literature review outlined (2013), there are important links between the constraints and opportunities at the individual, household and wider society and economy. It will take strong guidance from the program to combine to form a coherent narrative or set of narratives that bring these together.

Structural and Policy, Macro-Economic Issues

It will be critical that GrOW considers where it will place itself within broader policy debates and discussions on inclusive growth. This is where GrOW adds the most value. What is the role of macro-economic growth and how should it be promoted? Reiterating Kabeer's commissioned paper for the GrOW program, important to position the policy narratives well to contribute to the current dialogue on inclusive growth.

> "There is strong evidence that gender equality can promote economic growth… However, the converse relationship- that economic growth promotes gender equality – is less strong. Indeed, some of the fastest growing developing countries show the least signs of progress on basic gender equality outcomes. Formal regular waged work has the greatest transformative potential for women, but this potential has remained limited because of the lack of creation of decent jobs, and because of segmentation of labour markets.
>
> The paper suggests a research agenda that focus on constraints and choices that determine gendered patterns of labor market outcomes, both in terms of labor force participation as well as the segmented nature of the occupational structure. How do labour markets play out in different contexts and what are the precise barriers and blockages to women's mobility to better jobs or transition to higher value added enterprises? What changes are likely to ease the constraints on women's labor market options? And what forms of collective action around gender issues can drive positive change, at transnational, national and local levels? (Kabeer, 2012, p.3)

Kabeer touches on the importance of collective action. Evidence is also needed to address the structural barriers that persist in spite of modest gains. How GrOW positions its macro-economic growth narrative against or alongside other non-economic policies and interventions will be important. There is consensus that structural barriers persist and demand mutually reinforcing interventions:

> To date, interventions around WEE have focused primarily on equipping individual women with the skills and resources they need to compete in the market place – providing credit or training, for example, thus focusing on changing individual women rather than the context in which she is working. Now, however, there is widespread acceptance that WEE and gender quality will require measures to address the structural barriers that limit progress. It is clear that a body with the mandate of the High-Level Panel must look at the fundamental barriers to WEE if it is to achieve lasting progress, recognizing that change will require mutually reinforcing legal, social and economic interventions (Gender and Development Network, 2016, p. 6)

The UN Women Report (2016) captured a tension that was raised by several key external stakeholders as well. That is, it is not enough to assume that better integration of women into the labour market, even decent work, is enough. There is a fundamental need to structurally change the systems that are perpetuating the inequities which calls for harmonized economic and social policies. "Policy makers should move toward universal, rather than targeted transfers and services" (UN Women, 2016, p. 15). We need national laws and global conventions that are inclusive of the rights of poor and marginalized women and girls. Rights are indivisible and reinforcing of each other. They call for creating context-specific virtuous cycles of decent work, gender-responsive social protection and services alongside macro-economic policies that prioritize investment in human beings and the social objectives.

> Typically, the role of economic policies is seen primarily in terms of promoting economic growth, while social policies are supposed to address its "casualities" by redressing poverty and disadvantage and reducing inequality. But macroeconomic policies can pursue a broader set of goals, including gender equality and social justice. Conversely, well-designed social policies can enhance macroeconomic growth and post-crisis recovery through redistributive measures that increase employment, productivity and aggregate demand (UN Women, 2016, p.13).

These tensions and debates are not new and not new to the GrOW team. They are reiterated here to reinforce how important positioning will be. This evidence will have political implications globally. GrOW's program positioning related to structural change for WEE may, indeed, have an impact on global governance, macroeconomic and monetary policies.

Accountability that goes beyond national borders

In an increasingly integrated global economy, where states are outsourced, the realization of women's economic and social rights requires a wider framework of accountability, which encompasses the private sectors, States' actions outside their own borders and international organizations. These trends were reinforced by some of the external stakeholders. There is potential for some of the macro-economic research by GrOW to have transnational policy implications. This level of policy narrative may not be necessary but is worth considering as economic and related WEE constraints become increasingly globalized.

It is impossible to assume a "business as usual" approach to macro-economic and social policies at state levels when negative outcomes of WEE are tied to issues such as arms, drug and human trafficking, transnational phenomenon. "The current system of global governance exacerbates, rather than mitigates, the gender bias in macroeconomic policy." The Gender and Development Network (2016) urge reform of global governance institutions and more acknowledgement of the failure of economic policymakers to recognize that human rights agreements are application to their own sphere. They emphasize that WEE needs to be grounded in a sound analysis of corporate interests, land-grabs for resource development, the impact and exploitation by extractive industries and dangerous working conditions in global supply chains. Also, the precarious and often dangerous work of domestic, migrant, and factory workers that is often linked to these industries.

The African Civil Society report (2016) raises the critical issue of regional strategies and analysis to combat the rise in violence and its impact and interaction with WEE.

> In the last five years, for example, a number of African countries within the Sahel and the Horn of Africa have faced threats associated with extremist groups as well as organized crime, which include abductions, forced marriages, widespread violent attacks, arms, drug and human trafficking (Femnet, 2016).

The report argues that there needs to be compliance and enforcement to regional and international protocols such as the Maputo Protocol that links economic rights to social, legal and human rights in one framework. What are the conditions that allow countries to better comply with these protocols? The report notes that the main area of focus for WEE is in agriculture, business development, finance and Information Communication and Technology (ICT). These, as earlier stated, may be too micro and too focused on individual barriers to address structural constraints.

How to link the practical interventions and strategic structural barriers, therefore, is a critical issue for GrOW in its policy narratives? What role does national and international law and policy play? Where many of these conventions have been ratified, little has changed in practice. What brings “teeth” to these legal protocols and policies on the ground? Migrant workers and the globalized nature of work? What does it open and restrict for women? These are big urgent questions where strong evidence-based research has a role to play. Critical that GrOW participates in the broader dialogue of the most enduring questions.

3.1.3 GrOW Program Moderately well-positioned to have Strong Body of Evidence on Inclusive Growth

GrOW program is moderately well-positioned to have a strong body of evidence by the end of the program related to its original vision of “inclusive growth” in spite of having mixed quality of individual research projects. It will be necessary to be much more intentional and focused on the overall program-level body of evidence emerging to ensure that it is strong and speaks clearly to policy by the end of the program cycle. This section will focus on *what* is emerging regarding the scope of the evidence. The next section will examine how well positioned the program is to segment various policy stakeholders (*the who*).

The evaluation team judged the overall strength of the body of evidence against:

- Rigour standards based on individual research outputs, methodology and instruments for coherent theories of change, internal and external validity, reliability, cogency (DFID, 2014)
- Quality standards that situate rigour within gender contextualization, legitimacy and sectoral relevance (Ofir et al, 2016)
- Scope of evidence overall in terms of its ability to generate policy narratives across multiple settings (MSI International, 2012).

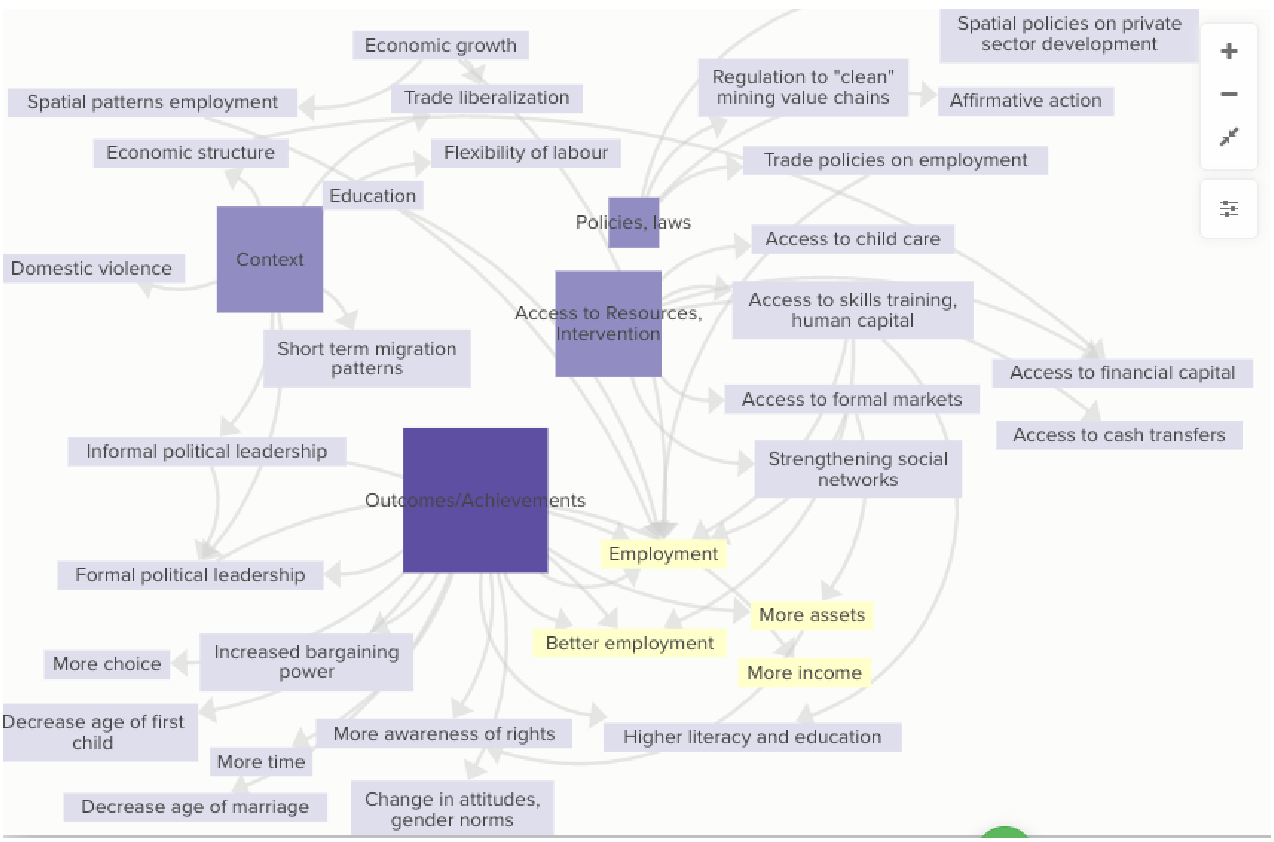
Early research outputs, including methodology and instruments, showed a mix of quality in the portfolio with 5 of the projects considered rigorous, relevant and well gender-contextualized. 5 could meet expectations with some corrective measures. 4 are rated weak in terms of positioning across the outcomes. Certainly, quality of research impacts not only the strength and scope of evidence but also policy uptake. It is important to note that ratings were done at research project level across all fourteen research projects but included an assessment of each early research output. Even where only methodology and instruments were provided, it was enough to assess the trajectory and whether the current approach was, indeed rigorous. As will be outlined later, public productivity, in other words the ability to share methodology, instruments and early outputs proved to be highly tied to quality including rigour.

The assessment of the body of evidence overall for GrOW takes into account not only individual project assessments of quality, but also project scope and the creation of a strong body of evidence across contexts. In order for the evidence to be strong enough to influence policy MSI International (2012) distinguishes between positive evidence from a few cases to evidence of impact from multiple settings and meta-analyses, where the body of evidence begins to be strong enough to develop a policy principle.

In this respect, the GrOW portfolio is stronger than the sum of its parts might appear. Since a few projects comprise consortiums and one of these was strongly rated with 15 sub-studies, there is considerable quality evidence coming out of the program. At the moment, however, the evidence feels piecemeal because it has mainly been understood at tproject levels, including project sub-studies and variables.

The following is a map of the variables being analyzed across the GrOW program by project level.

Figure 3.1 Evidence Mapping of GrOW Research Projects



The evidence was separated into context variables and conditions, policy foci, interventions and key positive and negative outcome variables. The yellow highlighted variables relate to narrowly understood economic variables while the other outcome variables at the bottom of the diagram show interactions with the economic variables in other spheres and forms of agency. It is helpful to see the various contextual or national features being examined such as economic structure, trade liberalization or spatial patterns of employment. In terms of interventions, the program examines access to a variety of resources including child care, skills training, cash transfers.

At the moment, the program level narratives are not entirely clear. In fairness, most of the focus, to date, has been on the projects. It will be important to move to focus on program level narratives since threads are certainly emerging that will have policy relevance both nationally and globally. It is helpful to be more specific about what is being fostered and scaled with respect to evidence which is critical to engaging the policy makers seriously.

We use a combination of the CMO (Context, Mechanisms [Policies; Interventions], Outcomes) systematic review framework and Kabeer's (2015) framework for Agency, Opportunities and Achievements to map out the emerging body of GrOW evidence around inclusive growth. These frameworks are helpful to distinguish economic interventions or access to resources (labour market, finance, training) from broader economic empowerment that would include agency (both intrinsic esteem and capacity as well as external norms that may constrain through stigma, discrimination or various forms of violence). Another distinction that is helpful is that between opportunities presented by interventions such as child care provision from policy opportunities such as trade policies on employment or affirmative action policies.

Certainly, the bulk of the body of quality evidence emerging examines the effects of macroeconomic growth policies on WEE or gender equity can have on growth. This macro level evidence will likely form a strong foundation for the policy narratives at the program level.

The following chart illustrates emerging policy narratives to provide the program with an understanding of how best to support filling the gaps going forward. Overall, the body of evidence can be brought under the narrative of inclusive growth.

Figure 3.2: Emerging Policy Narratives around Inclusive Growth

Three thematic themes and research questions	Emerging evidence and policy narratives	Observations
1. Barriers to WEE and to closing gender gaps in earning and productivity. How to overcome these barriers?	Patterns of sectoral and occupational segregation (macro-level data; South Africa including effectiveness of affirmative action policies; Colombia; more nuanced understanding of women's role in non-traditional sectors- artisanal mining sector) Micro-level RCT studies on gender gaps on assets and impacts on efficiency and productivity	Base of this evidence is quantitative macro-data and significant contribution from one project. Nevertheless, this thematic area is the most diverse.

	Barriers to skills training (distance) particularly for a mobility-constrained socio-cultural context like Pakistan Opportunities in the care economy. The relationship between paid and unpaid work, care responsibilities including child care. Possibility of contrasting community (Tanzania) vs. conditional cash transfer programs either as part of care economy or social policies. Positive pathways (informal to formal employment, education to work, unemployment to employment) and related factors (youth; gender; age of first birth; age of marriage) Bangladesh study complements Sub-Saharan African studies here.	Critical in the care economy work to have clear links to WEE as well as opportunities, not only dynamics. Also that household level dynamics can be nationally contextualized with time-use surveys or other quantitative data. Cash transfer programs add important element of highlighting social protection, social policies. Important that context-specific evidence across countries helps to highlight differences and nuances
2. How do specific patterns of economic growth and structural change affect WEE and gender quality?	Meta-analysis, systematic review, experimental research across hundreds of countries Impact of conflict on WEE in Northern Sri Lanka	Meta-level quantitative data largely from one project Important to get at asset entitlements, land, female headed household opportunities and constraints posed by conflict. Potential of paid and unpaid work studies (Nepal; Rwanda) to contribute to understanding of post-conflict contexts.
3. How does WEE and gender equality affect economic growth?	Theoretical models Do more equal countries grow faster? Short and long-term effects, inter-generational effects Gender inequities in growth (education, employment, pay) Main demographic and economic transmission channels	Meta-level quantitative data largely from one project Can contextual evidence be drawn on to support or reinforce this evidence?

These are elaborated further in the subsequent text in the next two sections.

Macro-Policies

In the policy arena, the main policies that will be addressed relate to macro-economic policies, growth, trade liberalization and employment strategies, which are certainly identified widely as a gap. GrOW is well-positioned to contribute to the effects of economic growth on WEE and WEE on economic growth. What does inclusive growth look like and how is it distinct from regular economic growth? The double direction of the analysis is quite new and warrants highlighting. There will be more understanding why sectoral and occupational segregation persists and where there has been leverage which will be a fundamental basis for future research. There will also be evidence related to “the feminization of labour,” the gendered relationship to informality and factors that affect movement from informal to formal employment. Under-examined sectors such as the artisanal mining sector and other male-dominated sectors will also provide new evidence. For most of these macro-economic issues, there is not only evidence at the country level but also across countries which will be critical to the global narratives around growth, globalization and gender. Critically, the micro-studies that support the inclusive growth narrative tend to be quantitative and cross-country as well and coming out of the same research project.

There is a lack of context specific nuanced analysis, for example with mixed methods, that help to contextualize some of the macro-trends. Unfortunately, one of the large multi-country research projects that set out do to this as well as link to social policies and infrastructure, is unlikely to largely fill this gap with quality outputs. However, some more focused analysis within this project may be possible.

The program level narrative needs to make some links between macro-economic policy and social policies since the harmonization of these came out clearly as a gap. Even if this evidence is thinner, the overall message should be clear. In a climate of increasing austerity and political populism emphasizing economic bottom lines, the program will want to position itself clearly with evidence knowing how it might be utilized. So even where social or public policy evidence is not the focus of this program overall, what does your body of evidence show?

The evidence on social or public policies is present but much more limited and some are still in question related to quality. Here, there is possibility of contributing to national policies related to child care, paid and unpaid work harmonization, education and skills training, less so to public policies related to infrastructure or tax, labour related policies and incentives. As stated earlier, growth alone may not address some of these issues. What else, beyond inclusive growth, is needed according to your evidence? Or are these examples of inclusive growth?

It is unfortunate that there is not a clearer policy narrative emerging related to the political participation research. Here it would have been interesting to understand how women’s informal political representation interacts with formal. Also, it would be important to understand how both forms of political representation influence policy priorities particularly those identified as gaps.

While some of the work had potential to get at social or public policies and infrastructure, particularly rural, the quality of this research is in question as well as its ability to get at these policy narratives. Much of the social policy-related research will stem from interventions and intervention research. Given how narrow these initiatives and programs often are, it will be essential that the research can contextualize the issue in the broader socio-economic and political context. More to the point, it will be critical that the policy makers engaged can be supported to contextualize it. So, for example, the child care studies touch on infrastructure as well as child care. The approach to policy uptake needs to go beyond the value of child care to see where the policy leverage is to position it in the broader need for enabling socio-economic policies and infrastructure. The same is true for paid and unpaid work, education or training-related policies to support transition to decent work.

Interventions

A large part of the GrOW portfolio, almost all of the first thematic area, is dedicated to programs or interventions such as child care, cash transfers, skills training, livelihood programs, economic empowerment programs. Particularly for the programmatic or practitioner-focused research, it will be important to ensure that projects move beyond the interventions as having intrinsic value and make clear outcome-related links to women's economic empowerment. It will also be important that the research findings do not focus too much on constraints but go on to identify leverage points or factors, conditions for providing counters to these trends. While barriers and constraints was identified as one of the research gaps, many of the barriers and constraints being explored have been well-documented. What is really required is some evidence on the other side of the thematic call, opportunities. For program coherence, it would be helpful to position these opportunities within the broader understanding of inclusive growth. For example, the research around child care, paid and unpaid work, care economy can demonstrate that growth on its own is not enough. There needs to be programming, policy and changed social norms that allow a more flexible and equitable gendered relationship to paid and unpaid work. More employment and more productivity will not necessarily change these dynamics. The thread of evidence emerging around the care economy and relationship between paid and unpaid work can certainly anchor the macro research.

Also, critical that the research moves well beyond specific program effectiveness to reliable findings that can be useful in other contexts. Even in the country where the research is taking place, gendered contextual analysis will be important to effectively ground and position the findings effectively for improved policy change beyond the program itself. Part of the focus on legitimacy and gendered contextualization was to ensure that this type of analysis was in place not only to design the methodological approach but to position it in policy narratives going forward.

Context

As both the external stakeholders and the reviews have emphasized, the WEE sector is lacking evidence about context-specific variations in WEE outcomes. Persistent WEE barriers in the labour market, in employment and decent work across contexts is quite well-documented as the earlier literature shows. The main gap is evidence that illustrates where, how, and under what conditions there has been leverage in addressing WEE barriers or why barriers persist in certain contexts. There is potential for GrOW to contribute here but the context-specific policy narratives are less clear partially because of quality issues with some of the research projects. Partially, the program selection did not select projects around a specific intervention or policy across a number of contexts. This, for example, was a suggestion repeated by two of the external stakeholders.

So, reviewing contextual gaps, one of the multi-country research projects has potential to shed light on the particular context-specific barriers and nuances of young women's relationship to marriage, education and work. At present, five of the six countries discuss factors related to age of first child birth and the relation to gendered engagement in education and labour market activities. One country is examining panel evidence on three pathways or transitions to work: school to work; informal to formal employment and unemployment to employment. The study in Bangladesh has strong potential to complement this Sub-Saharan African work. The context-specific (qualitative elements) in these studies will be important here to get at social norms, gender-based violence and the intersection of reproductive health issues, sexuality, work and education. There is a growing body of evidence describing the vicious cycle for girls and young women with limited economic opportunities and agency, including control over one's body, causing them to choose limited and unsafe economic "choices" such as early marriage, transactional sex and avoid school due to sexual harassment (ILO, 2016b). How the program deals overall with the question of "choice" and agency not only as personal choice but constrained by gendered social norms will be, in itself, an important contribution.

There is a large body of evidence describing the contradiction that the relative gains in education still combines with persistent inequality or very marginal gains in gender equity in labour force participation, particularly decent work (ILO, 2016; UN Women, 2016). There is not a lot of value-added to the policy dialogue to reinforce these constraints and barriers especially in trend data across countries. These are well-known and documented. More evidence is needed to understand why women are still locked in informal, longer hour and lower paid work, as well as the majority of unpaid family workers in spite of educational gains or transitions, trends in economic growth. GrOW can contribute if the evidence can help us to understand why gains in education have not led to a reduction in: gender gaps in the labour force; employment rates, particularly in income and access to decent work; and barriers to enable women to take full advantage of skills training. Also important is where the reverse trend or points of leverage exist and how policies can enhance their impact.

It would be important to understand what a conflict or post-conflict context means for women's economic empowerment and agency since there is conflicting evidence in this regard. Both with respect to the economic opportunities that may be opened up in conflict as well as what happens to entitlements and assets, during and post-conflict, particularly land. One of the research projects had potential to address this question but again quality and relevance issues may impede this contribution.

Outcomes

Outcomes can further be broken down into the areas of economic opportunities or decent work as well as how they interact with other outcomes (socio-economic, agency). A few projects are examining women's empowerment more broadly within which WEE sits. Such insights would be helpful to the sector both in terms of the outcome evidence and the way of measuring it. However, some of these projects are challenged with capturing these nuances effectively enough to have clear evidence by the end.

Agency is an important area in WEE and a difficult one to analyze and measure. The literature and stakeholders reinforce the persistence of these inequalities through, in part, gendered social norms. Critical to any research in WEE are insights as to how these social norms can be and are transformed or challenged over time. Agency is both internal (self-esteem, confidence, capacities, norm influences on choice) and external (discrimination, stigma and various forms of gender-based violence). It is yet, unclear, how much evidence the overall program will have related to agency. At best, some of the more focused projects may be able to capture these complex dynamics. Even a program-wide analysis or briefing related to the various ways that agency or social norms were analyzed would be a contribution to the growing dialogue around how we can understand these complexities better.

3.1.4 Grow Program positioning for policy uptake not entirely clear though clear threads are emerging

At the moment, the coherent Program-level policy narrative, or policy narratives more accurately, is not entirely clear though clear threads are emerging.

In terms of policy uptake, literature reviews and external stakeholders confirm that the inequalities that exist around WEE are not inevitable. Economic and social policies, both globally through conventions and nationally, have strong and proven potential to redress some of these negative outcomes. In order to discuss the potential for policy uptake given current positioning and trajectories it is helpful to distinguish the research in terms of audience levels. There seem to be four:

i. Strong policy narrative at the global level. E.g. Occupational and sectoral segregation.

ii. Evidence that provides a consistent narrative across contexts. E.g. Care economy

iii. Evidence that is more context specific and likely to be most relevant in-country. E.g. Cash transfers in Tanzania.

iv. Evidence that is not strong enough to publicize widely.

Global Level Policy Narratives

At the global level, there is a strong policy narrative developing related to WEE and economic growth, economic structures, trade liberalization, feminization of labour and WEE outcomes. These projects are highly academic and quantitative or meta-level for the most part. The systematic reviews focus on experimental research. As many respondents confirmed there can be a tension between scientific research of this nature and early policy engagement. As many respondents confirmed academically-focused research tends to see policy uptake as something that is not done until after the research is finished.

> *It is two different ways of looking at the world. Not surprising that it is so hard (policy uptake). In Academic research, you are protecting your discipline. Doesn't pay off in the academic world. Structural issue in academia that don't facilitate multi-disciplinary.*
>
> *We are really green when it comes to this side of things, policy uptake, communicating the results to policy makers.*
>
> Source: Project researchers

In some respects, early engagement is less an issue with this type of research particularly given that projects involved are well-connected to these agencies including the World Bank, IMF, IFC, various UN agencies, Wider University, national aid agencies. Where engagement and contextualization become important, is in questioning the occurrence of variations.

At a program level, depending on combined evidence by the end, some of the overall arguments will be important. Is it that economic growth is the tide that is "raising all ships" or is the economic growth and perhaps wider tax base permitting investments in social infrastructure, services? The reverse question is also possible for the work related to education and work transitions and pathways. Is an investment in education a good investment in economic growth?

> ***Education for Economic Growth:***
>
> While spending on education, health or water and sanitation is often seen as consumption, it can actually raise productivity, encourage private investment and stimulate higher rates of growth that can generate the taxes needed to pay back the debt. There are strong grounds for using deficit spending to finance social protection and basic social services, since critical investments in human capacities can ultimately create stronger economies and fairer societies.
>
> Source: UN Women, 2016, p.16.

UN Women makes this argument. The political participation work also had potential to make global claims relating women's political empowerment (both formal and informal) with women's economic empowerment. It does not seem well-positioned to do so given its current trajectory and the information available.

Evidence Across Contexts

There are two to three policy narratives or threads of evidence emerging across contexts that show the most promise include: the care economy and how best to harmonize paid and unpaid care work; women's role in small-scale and artisanal mining and other non-traditional /stereotypical sectors; and education to work pathways or transitions. Depending on whether there are adequate opportunities and leverage points identified related to the pathways, the program may provide more context-specific evidence to explain the "feminization of labour" or persistent informality of women's labour. How coherent a narrative can be derived from these threads will, in large part, depend on the quality of the individual projects by the end and their ability to position their arguments and policy messages in a harmonized way.

A key policy narrative developing for GrOW across contexts is to provide further evidence to explain why gains in access to education have not led to subsequent gains in labour force, employment gaps and quality of work in a country by country basis. There is not a lot of value-added to the policy dialogue to reinforce these constraints and barriers especially in trend data across countries. These are well-known and documented in the sector. What is helpful is understanding the differences and, critically, helping national policy makers to appreciate what is required in their context. One of the in-country researchers, a professor currently at Harvard reinforced the importance of regional foci and what it will take to ensure that the research has impact on the ground. "I have found regional or cross-country studies invaluable at a time where many African researchers tend towards single country studies. Language sometimes acts as a barrier in accessing literature for all team members. Bi-lingualism is a long-term objective."

The same is true for the projects related to the care economy and the transitions between the informal and formal economy. There is an important tension between supporting and validating unpaid care and informal work, improving its quality, and stimulating pathways to more formal work. Literature has shown that women often choose routes of informal work, finance and unpaid care work for highly legitimate reasons. Again, questions of agency and "choice" are important. What is required to address existing barriers is a highly contextual question. The gendered contextual analysis that was part of the assessment of legitimacy was not only to understand how well prepared the methodologies were. It was also to ensure that the research is well-contextualized to determine the most strategic paths to policy influence. Another element to the contextual research and its ability to defend its quality is the assurance that there is not an underlying bias toward any pathway or transition, such as formality. Rather, the evidence is uncovering leverage for increased opportunities for women and WEE.

More evidence is needed to understand why, in particular contexts, women are still locked in informal, longer hour and lower paid work, as well as the majority of unpaid family workers in spite of educational gains or transitions, trends in economic growth. The WEE sector requires more evidence that illustrates where, how, and under what conditions there has been leverage in addressing these barriers as well as imbalances in paid and unpaid work. Much of the literature points to national legislative and policy levers to address some of these barriers with the exception of affordable child care provision. Some of the research is well-positioned to influence policy makers nationally or regionally on issues of paid and unpaid work, child care, labour market transitions, skills training, how to support women in small-scale mining or extractives. Much will depend on if, by the end of the program, the evidence is reliable and relevant enough to be brought to regional or global levels to address structural barriers.

These pathways are more consistent and convincing across more than one context or population policy narratives and when they are outwardly-focused. Strong projects demonstrated that policy makers can not only influence research in constructive ways but early engagement with them can help to ensure that research is well positioned to influence without compromising ethics or standards of quality. It is worthwhile for the program as a whole to work to strengthen its policy coherence across several research projects addressing key outcome gaps identified in the WEE sector.

The current attention to, dialogue around and support for policy-oriented outputs is also helpful. It is evident that there are differing ideas and understanding of policy briefs. In one project, for example, though two policy briefs had been completed, the team agreed that one of these was more of briefing note on the methodology used than a policy brief. Policy briefs will need to make clearer links to policy audiences and messages. Policy briefs should be used strategically and sparingly with solid evidence to back.

3.1.5 Capacity building of in-country Researchers appears stronger than documentation is capturing

The GrOW program seems to be well-positioned to have built in-country researcher capacity, both at individual and institutional levels, better than is captured in the documentation. Capacity building was the most difficult outcome to assess. It is evident that capacity building is wide-spread and occurring. There were many different forms of doing so. Governance structure seemed to be the main determinant of the type of capacity building. How deep that capacity building went, however, was somewhat difficult to determine.

Capacity building is clearly taking place and seems quite pervasive across the program though it also seems to vary widely. The governance structure seemed the most reliable predictor of how capacity building was taking shape and even, to some extent, the strength of it, as noted by:

- **Mixed North-South consortiums** (Consortium refers to multi-country and mix of practitioners, academe);
- **Mixed North-South partnerships** (Two institutions with mix of practitioners and academe, North and South);
- **Academic consortiums or partnerships** (Largely academic, North and South; well-known universities); and
- **South-based institutions and partnerships** (Local research institutions).

Mixed North-South consortiums seemed to be the most administratively challenging structure given the diversity in locale, academic backgrounds and mix of practitioner and academe. There is complexity in ensuring coherence with this level of diversity. They seemed to have neither the efficiency of the tight 2-way partnerships, nor the consistent mode of operation and backstopping found in the largely academic consortiums and partnerships. These were also the most challenged in terms of quality. There was mixed reporting in terms of built capacity here. In one case, one of the in-country researchers claimed that very little capacity was built beyond their original skills. In other cases, in-country researchers discussed the ways in which the project has enriched their skills usually around a particular skill such as quantitative research skills. Also, usually the capacity building mainly took the form of the annual workshops when everyone was brought together.

The other governance structure that seemed to struggle with respect to capacity building was the **South-based institutions or partnerships**. Here there certainly could have been opportunities for capacity building across South-based partners. However, usually the partners were highly localized and shared similar backgrounds or understanding of the context. There was no external partner with which very different yet complementary skills and perspectives were brought. These structures, too, had issues with quality and coherence that seemed to be compounded by the team structure. Even where teams had a rich level of experience, some were overwhelmed in terms of administrative responsibilities making capacity building difficult.

The two governance structures that had consistently good examples and reports of capacity building were the more focused **Mixed North-South partnerships** where there was not only different skills and capacities brought together but there was a reported mutuality in respect for what each partner brought. The awareness was built in both directions. One solid example was where the PIs intentionally carved out mentoring gave additional responsibilities to up and coming researchers including primary authorship.

> Being deliberate in including junior researchers and giving them specific roles which they can take on and grow. As PIs we can take the lead in everything. It is easier. However, we let others take leads in certain places…..Yes, especially as first author. We provide inputs but they really needed to take the lead. Giving young researchers the space to grow and develop is key.

She also spoke well about the fact that a focus on capacity building is not necessarily exclusive of a focus on rigour. In fact, the two can reinforce one another.

> Unless we have strong evidence to inform programs we won't have impact. There is a need for evidence-based policy research. If we don't have the capacity we will fall short. Especially for Africa. A lot of the people who are trained abroad are at the end of their careers. Those poised to take position have not received the requisite training. We have to be cognizant of who will carry it forward. This shows the shortcoming of focusing on only now. We have to invest in building capacity of junior researchers to drive the next phase of research. (Of the three outcomes for the GrOW program) They are all important. They are tied and interrelated.

This PI also emphasized that their own capacity was built by the North-based university and mentoring in particular but that PI emphasized what they had also learned about tailoring to context from the South-based think-tank. Both PIs stressed the importance of a mutual respect for different skills sets and contributions.

In terms of the **academic consortiums**, there is only one with enough documentation and response to surveys and questions, to assess positioning for capacity. In both cases, however, the focus of capacity building was to bring in younger and newer graduate students (both Masters, Phds and to a lesser extent, post-docs). One respondent described this approach to a "graduate student factory," implying that they were providing more than perhaps receiving. One Project Officer expressed concern too in these projects that have too much reliance on research assistants. Those projects that rely on research assistants may not be building capacity as well as other projects that have built mentoring and learning opportunities into the design of the governance and research. However, interviews with these researchers found that the capacity focus was on providing them with a rich and applied experience for their

Challenge to Notion of North and South:

> Let me add something to be perfectly clear. We think of that university as north and those two as south. In terms of quality, they may be better than us. Capacity building goes in both directions. The best output would be that the links that these relationships survive the end of the project. We know we visit and can collaborate in the future.

Source: In-country researcher

research. The most important element of this experience was the cross-dialogue between countries and meta-analysis of why there were differences across contexts. One of these in-country researchers also challenged our notions of North and South as the earlier PI had done (See text box). This approach to capacity building does not necessarily address power structures in research. That is, bring those to the research table who otherwise may not have had access and give them really substantive opportunities to write, present and build their careers. There is much more emphasis on the principal investigators. The institutional partners tend to be well-established universities with established lead researchers. Using existing graduate students is already often a privileged group in these countries. In-country research was limited to relatively better-off countries, institutions and potentially researchers. However, this is still a legitimate form of capacity building and, as part of the portfolio, proves valid in combination with the other outcomes. It might be questionable to have this be the majority of the portfolio.

What is the role of research in redressing power dynamics? This was one of the questions raised during the Mid-Term Evaluation. Kabeer shared, and internal team dialogue reinforced, cautions that the move toward big data and experimental research in the name of rigour does not focus on economists and established universities and researchers.

Importance of Ample Time and Resources:

If you have a project that is three or four years, there is time. The funding that I get doesn't give me scope to do it very well. If I want to spend time helping them re-write, that takes time. So I think there is a real tension between wanting excellence and wanting capacity building. The tension is not so much with policy uptake. The capacity building.. it takes time and resources. I don't know how you are monitoring this? I don't think the RCT folks are interested in capacity building… They want a full time post-doc and it is their property. They seem driven by the opposite of collaboration. We carried out two qualitative evaluations of very similar proximity to these. I have a ground level for what the RCTs didn't consider-- the social identity, Muslims, Tribals. There should be more humility….the anthropology is not there. RCT have to learn to pick up anthropologists otherwise…I'm worried about this trend…

Source: Kabeer, 2015.

As many, including in-country researchers emphasized, building capacity well takes time, intentional research process design. It is critical, given the power dynamics and gender norms at play in this area that the questions and researchers continue to be diverse in all of the backgrounds that intersect with gender. Some researchers also pointed out that this is a long-term process. A program such as this is an important step in building the relationships and providing researchers and institutions with opportunities. Kabeer mentions the importance of providing ample time and resources for genuine capacity building. She did not dismiss the value of RCTs but cautioned the way they are done and combined (or not) with other more contextual analysis.

An in-country researcher confirms the time required:

> *There is a tension and one should be careful not to go in one direction. Then I would say has to have this time dimension very well in mind and act accordingly. The capacity building element doesn't work out short term or immediately. This tension could be resolved it results should not be expected too soon or process abandoned because we don't see results too soon. So going in the other direction, I do sympathize that high quality research should be the main objective. There is a tension being time spent on research and capacity building. So we shouldn't think do a lot of capacity building for one or two years and quality research follows. Long term process. Don't go too far in either of the two.*

Source: In-country researcher

3.2 Research Project Positioning for Quality, Uptake, Built in-country researcher Capacity

3.2.1 *Diamond map assessments: Research Projects show Mixed Quality*

For each research project, including the sub-studies, a diamond map assessment was done. These can be found in Attachment A. The evaluation team reviewed early research outputs and interviewed researchers to understand positioning for policy uptake and built capacity. The detailed assessments resulted in four groupings:

i. Strong across outcomes;

ii. Academically strong;

iii. Could improve and meet expectations with corrective measures; and

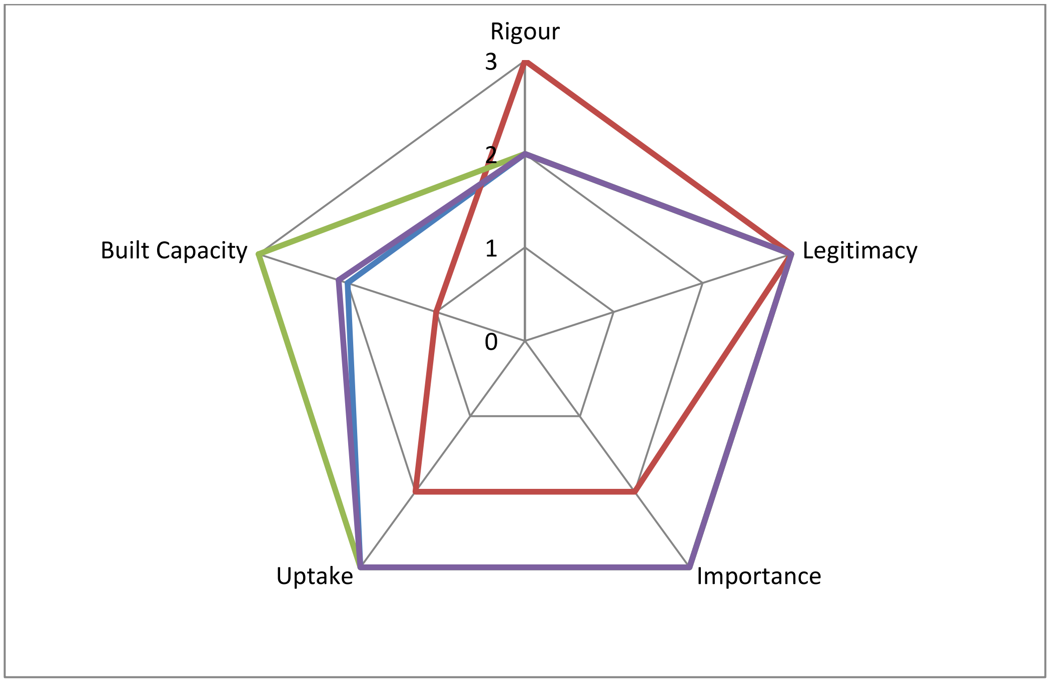
iv. Weak across outcomes.

Based on early research outputs, project assessment indicate a real mix of quality with 5 of the projects considered rigorous, relevant and well-contextualized, 4 that do not demonstrate high quality at this point and 5 projects that may meet expectations with some corrective measures.

3.2.2 *Strong across outcomes research projects*

There emerged, in the groupings, two types of strong research projects, those that were strong across all three outcomes (quality evidence; policy uptake and capacity building) and those that were academically strong. This section will deal with the former.

Figure 3.2 Diamond Maps for Projects Strong Across Outcomes



Strongly rated projects demonstrated a number of shared characteristics:

- A means for ensuring strong **coherence between qualitative and quantitative methods** whether through a strong principal investigator with mixed methods background or team structure that allowed ample iteration if analyses was separate;
- **Clarity on policy uptake** and early and continual engagement with key stakeholders for policy uptake; and
- **Good governance.** Complementary partnerships with at least two institutions, a mix of scientific rigour and practitioner or context knowledge, with at least one partner with a track-record for evidence-based policy research.

Rigour

Across the projects, high-quality research showed not only strong adherence to principles of rigour but they had a very strong coherence between quantitative and qualitative methods. The research project was able to both demonstrate cause and effect or trends (i.e. what is happening) usually at macro levels with qualitative, observational and participatory methods aimed at exploring why or how something is happening and how it might take shape in that particular context. In two cases, there was a mix of experimental research, and qualitative methods. These projects tended to be highly focused on a specific intervention or area. In the case of these projects, child care, skills training and women's role in the artisanal small-scale mining sector.

The following is an example where the mixed methods design is excellent and rigorous. The qualitative analysis is well-structured to help understand the quantitative trends.

Table 3.1 Example of a Strong Mixed Methods Design

DESIGN/PHASE	OBJECTIVE
Quantitative: Baseline surveys and Pilot experiments	Measure demand for skills (willingness) Measure enrollment into skills training
Qualitative: Fieldwork for Experimental design	Identification of factors driving reduced uptake with distance
Quantitative: Randomized Experiment to Understand Access Constraints for Women	What explains the distance effect? What can be done to mitigate it; to what extent those interventions help?
Qualitative: Ex-post qualitative research	Assessing SFM's impact on women's empowerment Understanding unexpected quant results

Objectives of all quantitative and qualitative analysis should be clear. It is helpful to see the level of detail and focus that allowed the research to use qualitative analysis to explain quantitative findings and factors. Also, interesting how the team has layered and phased their qualitative and quantitative analysis. The research will attempt to get at nuanced questions of economic empowerment as a process of building capacities and exercising agency over economic and social participation. Their analysis will include an analysis of persistent constraints distinguishing structural (social norms, opportunities) from intrinsic (personal motivations and aspirations). Given the design and focus, this research along with others in this grouping, may be able to provide some evidence about agency and social norms. The example of this research helps to illustrate why post field work it is difficult and cumbersome to attempt coherence where quantitative and qualitative analysis has occurred separately but was not designed together.

Clarity and Engagement for Policy Uptake

The last element of this research project worth noting is their approach to iterative policy engagement to maximize positioning for uptake. One of the PIs described the process of engaging the in-country policy makers, the academics involved, various stakeholders including private sector. Since the research showed high demand for skills training, the private sector stakeholders (skill training providers) were not interested in the why of low uptake. The researchers convinced them that this was important and how they are related. The government body (Skills Development Board) were interested in employable skills quite strictly and did not understand all of the focus on the household or a broader understanding of market failure in which they were situated. With repeated discussion and iteration, they were able to convince all of the stakeholders about the value of the research in its more nuanced form. The team will be documenting the whole process of arriving at evidence-based policy research for a Kennedy School of Business Case-Study. This project was highlighted to emphasize the importance of the policy uptake process as an iterative process of dialogue and negotiation rather than dialogue and dissemination of a completed research output. The former is much messier and resource-intensive meaning the policy makers and stakeholders involved must be few and strategic.

Perhaps because the research was so focused in all of the groupings, the research teams had access to and indeed, some were embedded in, policy fora. This allowed for early and continuous engagement with policy makers who could not lead the research but, in iteration with researchers, ensure that it was positioned in language and messaging.

Legitimacy and Contextualization

Similarly, projects within this group had very clear means of ensuring legitimacy and gendered contextualization of the research, not only through ethics approval and number of women researchers but through a carefully designed process of engagement. For some, contextualization was assured by who was brought to the research in the first place. For others, it was ensured in the process. In one project, they described feminist research where part of the approach to research is ongoing dialogue, reflection and iteration to uncover perspectives, norms, worldviews, is part what the research intends to achieve.

Other projects can learn from some of the mechanisms used to ensure legitimacy and gendered analysis:

- Strong emphasis on dialogue, iteration and reflection built into instruments and prompting;
- Strong gendered contextual analysis evident from positioning of research, adaptations, training guidelines for data collectors or researchers;
- Local male and female gender focal points form part of research process; gender paired data collection and analysis;
- High standards of ethics and transparency including mobile phone tips to ensure safety of participants and privacy of data;
- Areas for reflection and dialogue for the data collectors related to gendered analysis of the issues
- Translation of the research into local languages;
- Participatory methods that permit real sense-making and storytelling from participants;
- Respect for the safety, time, location of meeting, comfort level of participants evident from guidelines;
- Local stakeholder group that is drawn on continually to ensure legitimacy and contextual grounding of methodology, instruments, process

Good Governance

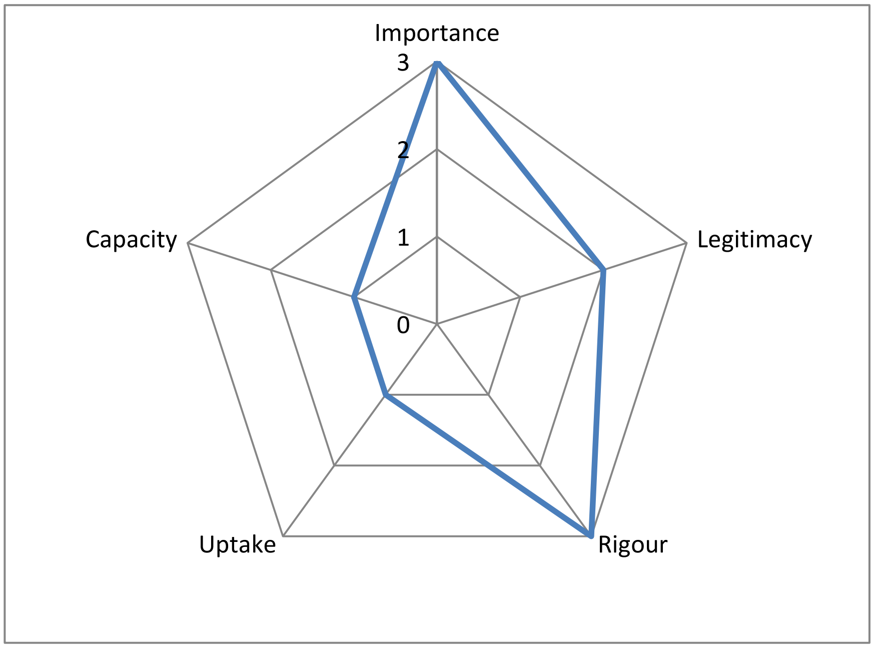
Both the governance structure and the project management seemed to play a large role in the quality of the project. Strong quality projects also had at least two research institutions with complementary skills involved including one with an established track-record for evidence-based policy research. Since coherence was so important to the quality of the project, mixed-methods experience proved important. In some cases, this was achieved by one or more lead researchers with this background ensuring coherence. In other cases, there were different qualitative and quantitative researchers and analyses but the process of iteration was regular enough and the guidelines and overarching design ensured consistency and coherence. The projects had a governance structure that ensured both substantive iteration, oversight and administrative capacity to handle the grant (and sub-studies where applicable) and capacity building.

In this grouping, evidently strong capacity building was a by-product of good governance. Capacity building was intentional and built on an already-strong structure of complementarity between research institution partners. It is even better where they have a previous history working together but that, as one project showed, is not enough to guarantee success or quality. There was an evident mutuality in the partnership that was strengthened largely by peer mentoring. The strong example of capacity building already outlined earlier, ensured intentional provision of opportunities for other researchers to take the lead in analysis or as lead other, with mentoring support and significant iteration and backstopping.

3.2.3 Academically strong research projects

Only one research project falls into this category though others could if the quality improves by the end of the Program. There was one more, arguably two more, academically-focused consortiums and an academically-focused partnership.

Figure 3.3 Academically Strong Research Project



The research projects that were academically strong, or in a position to be, were characterized by solid experienced researchers in their field. This GrOW research was an opportunity to expand and deepen the body of evidence that they had already significantly built.

Here the emphasis is on rigour and a scientific style of writing that is geared to peer-reviewed journals and high-level global audiences. There is, comparatively, less emphasis given to context-specific policy uptake, at least early engagement and ongoing dialogue, and to intensive or intentional capacity building as in the earlier grouping. Policy uptake is high level and strategic focused on policy makers that would understand and appreciate the nature of academic results with little tailoring required. Research topics have been chosen for their policy influence. Capacity building is the opportunity of Post-docs, PhD and Masters students to be involved in a multi-country research project.

One in-country researcher talked about the opportunity to work in such applied research. Usually, this diversity and links to field work is not possible in these PhD studies.

> *Really excited never worked in such a large project before. Interesting to be in such a diverse group. …Different universities. …Different countries. Rich set of different points of view. Meet and interact very often. Yearly workshop. Met for four days. Everyone presented their work. Amazing*

to see what everyone else is working on. So much expertise from the different countries. On the one hand, they have such detailed knowledge of the context. As a group we can have an overview of everything. Birdseye view. How are the contexts similar and different. That discussion could only happen in such a diverse research group. I just started my research. 20 or 30 people who know a lot more than I do. They know because they live.

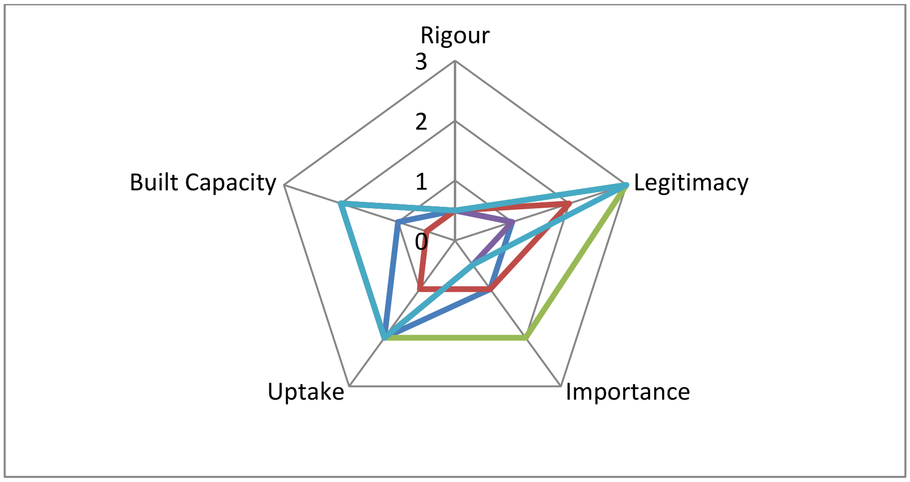
While this form of capacity building is certainly different than some of the other forms, particularly for non-academic in-country researchers, this form of capacity still seems to play an important role. This kind of research and research project is driven by the solid reputation and body of research of the principal investigators. Something important to consider related to legitimacy is that, reputation and strength of principal investigators is not enough to ensure quality. In four other research projects, with well-known and reputed principal investigators from highly renowned universities or research institutions struggled to ensure coherence and quality in their project and sub-studies. Their solid research background was not enough to ensure quality of the project. Regardless of investigator or institution's reputation, quality can never be assumed. There were questions raised of some of this research related to contextualization and nuance.

There is certainly a role for this type of research in an overall portfolio even if, as earlier described, it may not be advisable to have it dominate the portfolio. There is a strong place for experimental evidence and and big data but it does draw on data that can be of mixed quality as in the Demographic Household Surveys and also addresses very specific questions. It would be a worrying trend if this type of research were not accompanied with solid contextual mixed method or qualitative research in the broader portfolio.

3.2.4 Could improve meet expectations with corrective measures

Five research projects fell into the category of not currently meeting the "expect to see" in more than one parameter but could do so with corrective measures. It should be noted that one research project in this category is a bit of an outlier. It did not score highly enough to be in the category "strong across outcomes." However, it did not receive a "below expect to see" rating either. It has been placed in this category since it could benefit from corrective measure that would allow it to dramatically improve its positioning for relevance.

Figure 3.4 Projects that Could Meet Expectations with Corrective Measures



These research projects were rated in at least one key area as below expectations. There is a wide variation in where the rating was high and low. However, there were common elements. With corrective measures, there is reasonable hope that the project can meet expectations. In some cases, this may also mean adjusting expectations related to output deliverables.

There were some common elements present in this grouping as well:

- More practitioner-focused for the most part and issues of internal validity (selection bias or sampling);
- Issues of coherence between macro and micro analysis, clarity in the theory of change; and
- Issues of clarity and consequently, relevance.

Rigour for this grouping was below "expect to see" in one or more domains (with the exception of the one project) as the diagram shows. There were issues particularly around selection, sampling and controls. At times, the scientific approach was not defendable, the universe was not clear and the sampling was more based on partnerships, feasibility and logistics over what made sense for representation and internal validity. A caution or perhaps tension with these research projects that are embedded in practice or program evaluation is that their experience about what is necessary, while valuable, presumes or precludes the research. Research projects close to practice run the risk of being selection biased and leading analysis or instrument design based on pre-determined desired outcomes. This can lead to biases or leading questions in the instruments. For external validity, the research needs to move beyond programmatic effectiveness and that particular country to relevance beyond.

In some cases, the issue had more to do with clarity and coherence of research design. A common problem was a complicated theory of change that was trying to analyze too many conditions and variables to meaningfully get results. This led to weak methodological design and coherence between macro and micro analysis, quantitative and qualitative data. Often there was simply not enough focus and researchers could not easily explain the objectives of the different types analyses. In a few cases, qualitative and quantitative analysis were done too separately or without close enough iteration particularly in design.

Rigour and relevance were very closely tied in this grouping. Where coherence was lacking, it was difficult to determine the relevance to the WEE sector. In some cases, there was a straying from the original proposal and design that was more tightly coherent and relevant.

Project oversight for this grouping is mixed. In almost all cases, the POs increased the frequency of communication, a helpful lever. However, in some cases, it was difficult to understand the entire situation as there was not much formal documentation showing analysis of problems, potential strategies to address them and a timeline of follow up for accountability. There was often a long string of email communication but not a capture in a PMR, for example.

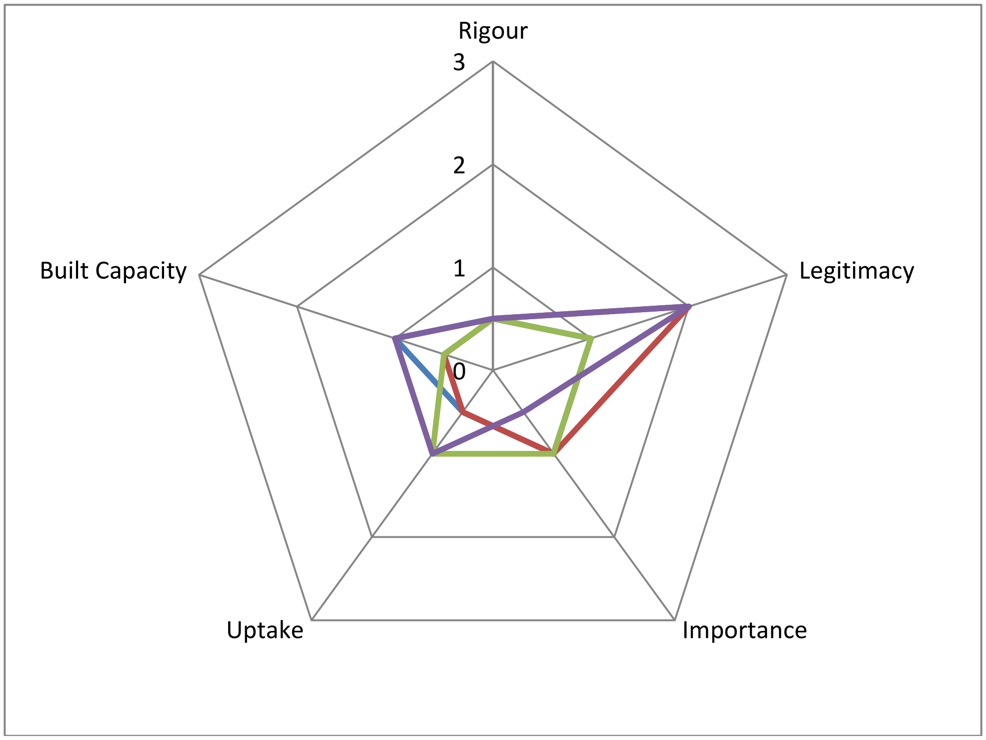
Another consideration for this grouping is the relationship between outputs and quality. In one case, the research project has struggled to meet deadlines and quality and has stayed to their intended research outputs, which are many. A key in-country researcher on the project shared in the interview that they are concerned about their ability to have coherence between quantitative and qualitative analysis. Yet, the team has just produced a number of papers. There seems to be a focus on “getting the work out” over taking the time to really ensure that the pieces are harmonized with solid coherence and consistency.

For this grouping, corrective mechanisms relate to identification of “pieces” within the broader research project that demonstrate quality. Remembering that ratings were done at the meta research project level, in some cases, there were individual outputs that were of strong quality. For example, a quantitative research paper may have been considered rigourous but had a lack of coherence with the qualitative paper so rated low overall for rigour. Going forward, it will be critical to identify outputs that can stand alone and be linked with broader evidence threads and policy narratives. In some cases, where sampling or field work was in question related to rigour, it must be considered if it is still possible to broaden the samples or ensure that they are representative by going back to the field. If that is no longer possible, these projects may not be able to meet expectations.

3.2.5 Research projects that are not well-positioned to meet outcomes

These four research projects are, given early outputs, not well positioned to meet outcomes.

Figure 3.5 Research Projects not Well-Positioned to meet outcomes



In these cases, not only are the projects not well-positioned to meet outcomes in more than one domain but the structure of the teams or the methodology has compounded poor quality. Given how far these projects are along in their process, it may be too late to change the structure. In most cases, the methodology has been carried out and field work has been conducted. Their shared characteristics:

- Weak clarity coherence (qualitative and quantitative often done separately);
- Single institutions without external expertise to complement or support; and
- Budget relative to their organization or other work (if low, less commitment; if too high hard to manage (single institution research projects).

These projects share a lack of clarity or focus on the overall theory of change or in harmonizing the qualitative and quantitative analysis. In one case, there is a lack of coherence with the sub-studies and to the overall WEE sector. In one case, the qualitative and quantitative analysis was done separately and have drifted from the WEE agenda. As a result, both lack quality and relevance as they now stand as early outputs. In two of the cases, the theory of change is not clear and not well matched to the methodology. It is difficult to see how reliable and meaningful results can be derived based on design and early outputs including instruments. In both of these projects, the focus is more broadly on women's empowerment and while that would be relevant and important in its relationship to WEE, they have not succeeded in capturing the nuance in a way that is both valid and reliable.

In all cases, the issues of quality in early outputs are compounded with structural issues. In one case, there is a clear lack of understanding between the qualitative and quantitative lead researchers about the other analysis making coherence extremely challenging. There is no one within the project who has a mixed methods background. As a result, the outputs read as completely separate, and less meaningful than originally-intended, documents. In one case, the overall design does not have coherence yet partners and sub-studies have been designed and field work conducted. Therefore, it is too late to find coherence here. It remains to be seen if individual publications or studies will have rigour, legitimacy and relevance but given what has been provided in early outputs, it does not seem likely.

In another case, the research institute lacks research capacity particularly, as the Project Officer noted, in quantitative analysis. Though they are partnered with a local research institute it is not enough. They require expertise in-house. The projects that performed best had partners involved substantively in design and research. An advisory role may not be enough.

Finally, budgetary issues in terms of the relative size of the funding in the organization can play a significant role in two directions. The size of the funding can be large and overwhelm the administrative capacity of the institution, as a Project Officer said was the case in one of the projects. In another, the opposite problem may be true. A relatively small amount of funding is provided to a well-known university making it potentially small in the mix of other funding sources and projects. In this case, it is difficult to know what the issues are because the project is so non-responsive in their deliverables. Even for the MTE, the survey and follow-up questions were not completed or given a response of any kind.

3.2.6 Project management effectiveness varied and highly tied to quality and governance structure

Project management effectiveness, styles and governance structures were as varied as the research topics and highly tied to quality. There seemed to be two elements in highly-rated projects that were not apparent in weaker projects: coherence and consistency; productivity particularly public productivity. This was what we could see of stronger projects. The underlying levers behind these conditions seemed to be good governance enhanced by structure as well as a strong process:

- Governance structure demonstrating coherence and consistency; and
- Process for ensuring coherence, and consistency (usually through a mix of governance, how the team was structured and a lead or leads with a specific role of coherence).

There were a few different types of governance structures in the program as earlier described:

i. Mixed North-South consortiums (Consortium refers to multi-country and mix of practitioners, academe)
ii. Mixed North-South partnerships (Two institutions with mix of practitioners and academe, North and South)
iii. Academic consortiums or partnerships (Largely academic, North and South; well-known universities)
iv. South-based institutions and partnerships (Local research institutions)

As earlier described related to capacity building, implications for administrative capacity and substantive coherence, for the most part the high-rated projects, tended to be North-South partnerships or academic consortiums.

> *I think the trick is in having a balance. It shouldn't be either or kind of situation. I think a project like the one we have here that is an opportunity to build capacity as well as have rigour in whatever we come up with. The data collected or the papers that are written. The only thing is*

that the process is going to be longer. Those who are experienced have to have the will and the time to work with others. The way we are approaching it. The local teams write the initial draft. Then we turn it to the core team (PI local) and they raise questions and give comments and then to the main PI. There was sufficient time that was built into the project to allow that interaction.

Source: In-country Researcher

3.2.7 Several program-level levers identified that helped to ensure quality

There were several levers identified at the Program level related to ensure quality. There also seem to be key periods of interjection for the program to ensure projects and the program overall is on a strong path to achieve the quality-related outcomes:

- Design of calls for proposal for coherence
- Selection and rejection, risk identification
- Team structure and methodology design prior to field work
- Ongoing project oversight, support
- Accountability for quality evidence-based policy research

While levers related to policy uptake and capacity building are fewer, there are important elements related to making connections, outcome-focused monitoring. Going forward, the last two elements can be emphasized: oversight and support; accountability mechanisms.

There may be a need to drive the research a bit more in terms of emphasis and stronger mechanisms to ensure that it is quality and also feeding a relevant policy narrative.

Quality Control Mechanisms

Original call for Proposals

The original call for proposals and the connection between those themes certainly has an effect on the coherence of the program overall. It is not surprising that the majority of projects from the first call focus on barriers as that was part of the call for proposals. Unfortunately, over time, some of the projects have focused less on opportunities or leverage points and more on barriers. Also, having some of the meta-level questions such as the relationship between WEE and economic growth in the second and third rounds meant that the mixed methods and qualitative research from round one could not effectively build on trends coming out of meta-level, systematic reviews or cross-country analyses. It is understandable and, as explained, the process for thematic calls was somewhat organic and there is often a natural drift from proposal to practice. However, this could be a consideration going forward, particularly in future calls.

Selection and rejection, risk identification

Selection appears strong based on the criteria with some exceptions. One of the projects did not have coherence even at the time of selection and the coherence has not improved with time. In a couple of other cases, there were some questions about whether projects should have been selected given the weak research capacity identified combined with a lack of external partner to round out that capacity. One example was a lack of research capacity identified as a risk compounded with the size of the funding relative to the overall organization's budget and the fact that there was no external partner to complement capacities and handle the research logistically. These issues reinforced by the project assessment risk analysis in the Project Assessment Documents which were extremely detailed, helpful for oversight and quite accurate. These risks might be helpful considerations for future given that in nearly all of the cases where methodological risks were identified, the quality and performance of the project in their current trajectory has not been strong. The rationale for rejecting one of the originally approved projects is sound and an indication that the program has "teeth" when it comes to accountability.

Team Structure and Methodology Design Prior to Fieldwork

In those cases, where either team structure or methodological design issues were problematic early intervention was critical to make changes. For others who have already begun fieldwork and finalized instrument design, addressing some of the coherence issues are difficult at this stage in the process. In future, it would be helpful to make these elements of methodological rigour and structure for coherence conditional on continued funding to go to the field or continue the research.

Going forward, the only mechanism available to support projects with weak capacity coupled with weak governance, is to suggest or provide external advisory support such as has been provided in some cases. In some cases, given dynamics this may not work well or be timely given where the team is in the cycle.

Ongoing Project Oversight, Support

Principal Investigators shared the view that the level of accompaniment and support has been strong and welcome. While many felt that the reporting requirements are a little higher than average funding sources, overall support is appropriate. PIs particularly noted being connected to donor networks and policy fora as well as to resources and expertise across the program.

In some cases, where the project is not performing well, the project officer increased communication to monthly and the reporting to every three months and the communication became even more frequent. Early and continual corrective measures were important in project oversight to help research projects improve.

In terms of accountability, where project officers had continuous and well-documented corrective measures they have helped to improve project accountability to outcomes. The following is the type of documentation that proved important to accompany the support:

- I also re-iterated my concerns about the project's overall slow progress and the quality of the draft outputs received. We discussed the project's strategy for capacity building and quality control, and the role of the project's co-ordination team therein. I shared copies of the youth employment Working Papers we recently completed as an example of the kind of outputs we should be aiming at this stage in the project's cycle
- We discussed capacity needs across countries.
- Overall, the project is at a critical junction whereby its performance in the next six months will determine whether it will succeed or fail. As a first step, the project needs to complete…… The PIs agreed.

- We need to closely monitor progress as the project cannot afford any further slippage.

Source: GrOW Project Officers

This is the level of engagement that proved constructive from the POs. The PO was able to sort out with the team, what elements of the delays seem to have been due to external reasons (what elements related to "lack of follow-up and guidance" and where capacity supports would be necessary. The PO also ensured the PIs were agreed with their collective analysis and solutions going forward. Finally, there were specific time-lines for follow up and expectations of results or changes. This is the type of early corrective supports, as with provision of solid output examples, that are not only highly constructive but timely.

This research project has since improved a great deal even though they are still in the grouping that may meet expectations with corrective measures. This was not a research project that had risks flagged because some of the researchers had worked together prior. So, the unexpected can happen. What is important is that the Program could likely have done little more.

In other cases, the level of documentation is not tied to the performance of the research project. In one project, weak across outcomes and behind in its deliverables and performance there is not documentation past April 2016, nor copies of methodology, instruments or early outputs. While it is difficult to assess quality completely, lack of outputs at this stage in the process particularly instruments may be a red-flag to serious problems with quality and ability to deliver. In this case, there is evidence of frequent email correspondence stating expectations and standards for the program.

While there seems to be highly supportive accompaniment with all of the research projects, there is not detailed documentation of this accompaniment for most of the projects including those where there are quality issues. Of the 8 research projects that fall below "expect to see," including those that may still improve with corrective measures (3), only 2 have a formal Project Monitoring Report. It is understood that PMRs were mainly used as trip reports. However, it is worth considering where some form of monitoring is not necessary on a more continual basis to formalize what is happening informally with communication and back-stopping. This capture is particularly important when projects begin to fall behind or face quality or team dynamic issues.

Clarity and appreciation for what is meant by rigour and quality

Going forward it is important that there is both clarity and appreciation for differing views on rigour and quality. They do not have to be the same views to be mutually reinforcing. Differing views of the various donor representatives on the parameters and standards for "quality" led to delays in establishing standards. While all of the donor agencies on the Executive Committee are committed to the outcomes, there were varying levels of emphases particularly related to "rigour." The bigger question at hand is how the various donor agencies define quality evidence which is a tension and debate in the broader field of development too. On one hand, rigour is key. As one representative said "there is more and more pressure to show value for money. We won't take any sub-standard research. There is a very big push for "racheting" up. We place more emphasis on quality, rigour, credibility." The sense, shared by some in the internal dialogue, is that while all three outcomes are important, rigour is the critical pre-condition for policy uptake and capacity building.

Another view is that rigour, legitimacy, policy uptake and capacity building are all inter-related and form quality research. As another representative put it, "These three outcomes are inter-related and reinforce one another. We have an appetite for risk. We would rather give an opportunity to learn with and fail. We look as this as relationships not as a one off but an ongoing accompaniment. Power structures don't change magically over three years." These programs are not only about producing "rigorous" evidence but about building relationships for which this is only one step, even if something doesn't quite work. The importance of balancing the three outcomes was also confirmed by many in the internal dialogue, particularly in-country researchers who spoke to the few opportunities to have capacity built into the program design.

In fact, these contrasting but overlapping views on quality and rigour are more widely held in the field of development and within the program. They are not exclusive views, but overlapping. The key, and perhaps the art, is finding how these outcomes can effectively reinforce one another. While the assessments confirm that a minimum of rigour is necessary to have policy uptake and capacity building, it cannot be assumed that rigour leads to the other two outcomes. As many, including in-country researchers emphasized, building capacity well takes time, resources, intentional research process design. It is critical, given the power dynamics and gender norms at play in this area, that the questions and researchers continue to be diverse in all of the backgrounds that intersect with gender.

Accountability and standards for quality evidence-based policy

Building a strong body of quality evidence is one of the intended program outcomes. Accountability when projects are not on course to do so is critical. The conversations related to standard-setting are important and timely. The EC has made important strides in agreeing on standards and outputs. The program has now established standards for quality and rigour according to common principles (coherence; internal and external validity; reliability; cogency etc.) as well as parameters for the research outputs as found below.

GrOW Internal Discussion of Research Outputs

Working Paper: A working paper for an evaluation involves detailed reports of methodology, analysis, results (or results-in-progress) and policy implications. These papers are pre-publication, but should be of a sufficient standard that they can be used to inform policy. The GROW team centrally should assure they are of sufficient quality.

Synthesis Paper: Studies analyzing data from multiple reports in order to identify larger trends across a particular region and theme, with emphasis on drawing policy recommendations.

Synthesis papers should be peer reviewed (peer reviewers should be named).

Ideally a synthesis paper would also include some economic evaluations of interventions (both identifying where tackling a policy problem could provide a positive rate of return, and data on the cost of the intervention in question).

Policy Brief: Crisp and comprehensible summaries of evaluations and their findings, with a strong emphasis on the resulting practical policy advice.

The audience is a government official or another policy maker. The policy brief focuses on a single issue, sharing the key evidence that he or she needs to know, providing high quality analysis, proposing practical solutions and recommending improvements.

It is helpful to see how this is playing out in the PO supports. For example, this following is an excerpt from an email of a PO to a PI reminding them of the process:

i. Interim research processes: Principal Investigator ensures research process is consistent with research quality standards for rigour, legitimacy, importance and positioning for uptake.

ii. Grey literature: Principal Investigator approves release of documents listed as grey literature.

iii. Working Paper: All working papers submitted to GrOW need to be signed off by Principal Investigator. If the project research team has its own institutional working paper series, working paper has to comply with internal standards.

iv. Journal articles: All draft journal articles signed off by Principal Investigator and/or institutional editorial policy

v. Policy Briefs: GrOW provided a tutorial on briefs. GrOW provides template for briefs and quality checklist. All policy briefs produced (making use of tools provided) are reviewed through internal institutional processes of the project research team AND must pass checklist (stamp of approval from lead investigator).

What is evident is that there are still varying ideas and applications of these outputs. Working papers, analytical studies, synthesis papers have wide applications and are interchanged. In some cases, a synthesis paper brings various country level studies. In other cases, it is bringing together qualitative and quantitative analysis within a country. One observation is that some of these outputs include or imply the audience and some do not. Those that include the audience are clearer or perhaps a little less open to confusion or varied application. Synthesis papers, as originally envisioned, may not be practical or recommended given emerging threads of evidence and policy narratives. It is quite possible that there will not be enough evidence or enough quality evidence to effectively pull these synthesis papers together. Rather, a more effective use of the GrOW team's time will be to ensure that the forthcoming research outputs are strong evidence-based and policy directed that align to provide clear policy narratives.

It will be important to establish these standards and ensure that they are upheld. The program's legitimacy and even project evidence within it can be compromised where quality is called into question, particularly where the level of quality may not be strong enough to be reliable or relevant for policy or practice. Clarity on the parameters of the outputs is related and relevant. There are currently varying levels of quality outputs using similar terms for different outputs. It would strengthen the program coherence to be sure that they are being used consistently and to potentially build quality into the names to better support projects as well as hold them accountable.

At the moment, where quality of the project is an issue, there are varying mechanisms to deal with the situation including increased communication, provision of solid examples of outputs, visits to back-stop the process. The program has not used, for example, stronger mechanisms such as withholding funds, attempting to drive the research focus or links to policy more, denying or delaying rights to distribute the research in certain forms or venues unless quality is met. It is important to ask: what happens if some of these projects continue with issues of quality and coherence and nothing improves? What is the process going forward? Organizational culture came up where the focus is on accompaniment and support, building and establishing longer-term relationships even if the quality is not there at the moment. One point made was that the institutions have started at different points and it may be fairer to judge according to where the institutions and project began. Indeed, in-country researchers stressed the importance of this type of accompaniment as well.

The program was widely cited from researchers even Executive Committee members and external stakeholders as being supportive, one of learning and dialogue. Evident even in dialogue throughout the evaluation process is the willingness to learn and heavily debate what is working and what is not working, how to improve and adapt. Both the will, at a team level, and frameworks to get at the nuances are both critical. Good documentation can support this dual-focus of learning and accountability, differentiated supports particularly where the focus on outcomes.

While support is certainly the overarching culture and aim of the program, both accountability to quality and support can be provided if properly handled. For some who require support, it may be possible to focus more on capacity building and the longer-term relationship and outputs rather than pushing immediate published outputs of poor quality. For others, stronger accountability mechanisms may be required. It is important to note that quality issues exist for both North and South based institutions and researchers from the Global North and South.

Policy Uptake and Capacity Building

Policy uptake has been a priority for the program particularly early engagement. In fact, many identified the balance of outcomes as a key value-added not found in many research projects.

Nevertheless, the projects generally did not generally score well in policy uptake. In part, many of the researchers, given their backgrounds, see policy uptake as something that is done once the research is completed. So, in part, a product of timing. There seems to be some confusion between policy uptake, which requires prioritized engagement with very few influencers, and communication and dissemination which is much wider and varied across a number of stakeholders. While the program emphasized early engagement, the experience and attitudes of many of the researchers (who come from an academic rather than practitioner background) combined with monitoring, may have worked against the program strategy.

Multiple stakeholders agreed that GrOW has played an important role in making projects aware or connecting them to policy makers, forums and networks.

> As far as I know IDRC puts a great deal of emphasis of doing policy engagement from the start. You get interest with the kinds of people who might be interested. That is not academic practice. They about reputation. The career.
>
> PIs confirmed that the support that the GrOW Program provided in terms of links to networks, UN High Panel etc. have been very important and helpful.

Source: Principal investigators, Executive Committee members, few external stakeholders.

It became evident reviewing documentation and in interviews that many projects have collapsed their communication and dissemination strategy with their policy uptake strategy. They are different in important ways as the Program staff is aware. The monitoring reports seem to exacerbate the collapsing. It was very difficult at first to get a sense of their priorities for policy influence both in area and messaging as well as the particular policy makers and arenas based on a sound contextual analysis. Those policy makers would be much fewer and the strategies more tailored with this analysis. The majority of projects had quite comprehensive listings of various stakeholders. Some were able to name specific influencers and how their research policy narrative was evolving. However, largely, there was not a tailored prioritization or differentiated strategy for different policy audiences as defined in the assessment framework. Number of policy makers, forums and discussions are more outputs than good indicators for outcomes. It would have been helpful to have a few focused outcomes combined with a short narrative.

Capacity building was similar. As earlier described has been difficult to detect in terms of outcomes or depth of capacity. In part, the monitoring framework had the same issues with a focus on outputs, number of in-country researchers. While it was excellent to see the gender-disaggregation, without meaningful outcome related indicators or narrative focused on outcomes, it is possible that women in these positions could be token. That is a concern or at least a question in at least one project where women were brought in after the discrepancies were pointed out. The assessment framework had as a like to see, a baseline of assessment of competencies that could be tracked over time. Very few did anything so detailed. However, a few projects were able to list some of the key knowledge and skill areas developed in particular team members. In a program as complex as this, it is difficult to distinguish, without effective outcome measures, the difference between learning by doing as seemed to be the case in most circumstances, with a more intentional and differentiated approach to capacity building as outlined in the highly rated projects section.

4 Key Findings

4.1 Overall, the Program is moderately well-positioned to achieve its outcomes.

As assessment of both research projects and the overall GrOW Program have shown, GrOW is moderately well-positioned to achieve it outcomes. There is a strong body of evidence emerging that is linked to a number of policy narratives that fill key gaps in the WEE sector as identified by key WEE literature reviews and external stakeholders. Capacity building is widely evident, both for in-country researchers and institutions, according to interviews and assessment though least clear of the outcomes.

Based on early research outputs, project assessments indicate a real mix of quality with 5 of the projects considered rigorous, relevant and well-contextualized, 4 that do not demonstrate high quality at this point and 5 projects that may meet expectations with some corrective measures. Nevertheless, The GrOW portfolio is stronger than the sum of its parts might appear. There is moderately strong body of evidence emerging with strong policy threads evident. Leveraging these will be critical to Program-level influence on policy and practice.

This result shows the importance of organizing research projects in a programmatic structure where policy narratives gain more coherence and potential influence. The program is demonstrating that the three outcomes are achievable and compatible but cannot be assumed.

4.2 Strong threads of a body of evidence emerging

Finding 1: The GrOW Program is stronger than the sum of its parts might appear. Quality of individual research projects is mixed, but there are strong threads of a body of evidence emerging.

Based on early research outputs, project assessments indicate a real mix of quality with 5 of the projects considered rigorous, relevant and well-contextualized, 4 that do not demonstrate high quality at this point and 5 projects that may meet expectations with some corrective measures. Nevertheless, The GrOW portfolio is stronger than the sum of its parts might appear. There is moderately strong body of evidence emerging with strong policy threads evident. Leveraging these will be critical to Program-level influence on policy and practice.

Finding 2: Quality research projects had strong governance and coherence between quantitative and qualitative methods, as well as early positioning and clarity on policy strategy.

The overall program can learn from the elements of the highest rated projects. These included a strong coherence between qualitative and quantitative methods, clarity and specificity on the policy uptake strategy including early and continual engagement and good governance. The governance structure proved important not only composition but in ensuring coherence, consistency and complementarity between partners.

4.3 Policy narratives emerging that address key gaps

Finding 3: There are a few policy narratives emerging that address key gaps in the WEE sector. Coherence and clarity is key to adding value to the sector.

While at present the policy narrative at the program level is not entirely clear, there are some key policy narrative threads emerging with clear relevance to the WEE sector. Policy narratives emerging include: transforming work (sectoral and occupational segregation, harmonization of work and family responsibilities, pathways or transitions to work and decent work; role of women in non-traditional sectors such as mining) as well as links and factors affecting macro-economic growth and WEE. Globally, the main value-added relates to the link between WEE and macro-economic policies, growth, trade liberalization and employment strategies. Evidence related to social or public policies is much more limited especially because quality of evidence is in question with the research related to these areas.

Much will depend on if, by the end of the program, the body of evidence can make effective links between the macroeconomic trend evidence with the more nuanced and context specific evidence at country and micro levels. It will also be critical to begin understanding what policy messages are emerging related to inclusive growth or economic growth and WEE? How is inclusive growth different than regular economic growth? How do they relate and harmonize with social policies? What, if anything, is required to complement economic growth to ensure WEE? In what circumstance can greater WEE and gender equality also contribute to economic growth? How do these narratives change by context?

In terms of context-specific evidence, there is evidence across contexts such as those related to the care economy, paid and unpaid work and responsibilities and also transitions from education to employment, also from informal to formal employment or more decent work. This work around transitions and trade-offs will fill important gaps if it is coherent, reliable and relevant enough to be brought beyond specific contexts to address national and regional policy levers. Finally, there is evidence that is highly country and context specific and will play an important role in those national policy arenas. These include cash transfers, skills training, empowerment programs. It will be important that the evidences move beyond the focus on poor women's constraints or inherent value in empowerment programs, child care and help to link these to broader policy levers and contextual leverage points including for agency.

Finally, there is evidence that is not high enough quality to publicized widely or use to influence policy, at least not in its current trajectory.

4.4 Capacity building of in-country researchers and institutions appears stronger than documentation captures.

Finding 4: Capacity building varied and apparent, stronger than documentation captures

There is apparent capacity being built across the program though it varies widely and the depth of capacity being built is difficult to detect with current monitoring that is mainly focused on outputs. An intentional approach to capacity building such as mentoring and allowing researchers other than the PIs to take manageable leads was key.

The governance structure seems to play an important role in the projects' effectiveness at capacity building and improving mutual complementarities. Mixed North-South partnerships and academic consortiums or partnerships seemed to rate more highly in quality, coherence and capacity building in part due to the structure, administrative issues and complementarities.

As such, there is an important role that GrOW plays in building institutional capacity. Outcome-oriented indicators, particularly at the institutional level, and focus on what is most critical to meet the other outcomes will be helpful going forward. Some projects will require more support than others to meet expectations.

The three program outcomes are mutually reinforcing both positively and negatively. However, capacity building takes time and an intentional approach. Many report capacity building to be the main area neglected if there are issues of timeline or budget to ensure rigour.

There is apparent capacity being built across the program though it varies widely and the depth of capacity being built is difficult to detect with current monitoring that is mainly focused on outputs. An intentional approach to capacity building such as mentoring and allowing researchers other than the PIs to take manageable leads was key.

Finding 5: Governance structure was important for building capacity, as well as an intentional approach.

The governance structure seems to play an important role in the projects' effectiveness at capacity building and improving mutual complementarities. Mixed North-South partnerships and academic consortiums or partnerships seemed to rate more highly in quality, coherence and capacity building in part due to the structure, administrative issues and complementarities.

As such, there is an important role that GrOW plays in building institutional capacity. Outcome-oriented indicators, particularly at the institutional level, and focus on what is most critical to meet the other outcomes will be helpful going forward. Some projects will require more support than others to meet expectations.

The three program outcomes are mutually reinforcing both positively and negatively. However, capacity building takes time and an intentional approach. Many report capacity building to be the main area neglected if there are issues of timeline or budget to ensure rigour.

4.5 There are project and program levers that can likely support some, not all, weak research projects in coming to an acceptable level of quality, uptake and capacity building.

There are project and program levers that have been important in ensuring that the research projects meet their outcomes. At the project level, team and governance structure mattered to ensure complementary skills and manageable administration while still building capacity. Sound methodological design, particularly where qualitative and quantitative or experimental were combined, was critical. Essential skills in the team included solid mixed-methods research experience and evidence-based policy research and engagement. Stronger projects were effective at bringing in complementary skills to their internal team either through partners, advisors or an advisory committee. Good governance combined with a strong process of guidelines for constancy, standards, frequent dialogue and iteration were also key. Going forward, it is difficult to change governance structures but some projects may benefit from more external advisory support or stronger project oversight and direction particularly for rigour and policy relevance.

The program levers also helped ensure strong positioning for outcomes. Early in the program, call design, selection, rejection, risk identification were key. Another key intervention point is prior to finalizing instruments and methodology and beginning field work. Currently the discussions and finalization of standards is critical as are accountability mechanisms to ensure that they are followed. Standards, close accompaniment, especially for weaker projects, and outcome-focused documentation played a key role in ensuring quality research and policy uptake. These will be critical going forward. In fact, the program may reconsider the quantity of outputs expected against fewer quality research outputs that are more tailored to particular policy audiences. Early correct measures and clearly documented accountability were important to support projects that were struggling in different ways. Stronger mechanisms such as stronger research or output direction or withholding funds were not used but may be important going forward.

In some cases, quality issues in the research are compounded with the structure of governance and the field work is complete making corrective action challenging. While the research can be slightly tweaked, it may not be in a position to dramatically improve in the remaining time. These projects, however, will affect the legitimacy and reputation of the overall program so how they are handled will be critical. In these cases, it might be important for the reputation of the overall program to revisit the public sharing of these research outputs and focus, instead, on building capacity.

4.6 The overall program has clear value to the WEE sector. Emphasis from projects to program level positioning will be critical to ensure this relevance going forward.

GrOW’s strength and its challenge is the diversity of the program. The overall program has clear value to the WEE sector. It demonstrates in its diversity the importance of mixed researchers from the Global North and South, a mix of academic and practitioner research, mixed-methods and a mix of global meta-data, cross-country and country-specific research. They key to going forward will be to harness and leverage this diversity into coherent but nuanced policy narratives.

It is not entirely surprising that, until now and with the exception of the Mid-term workshop, the emphasis within GrOW has mainly been on the research projects, provision of support and accompaniment to ensure positioning for outcomes. The Mid-term workshop, external policy meetings have begun focus on the overall program positioning around inclusive growth. It will be critical that program focus is even more sharply tuned and communicated. Ensuring a program focus going forward will be critical to leverage the value-added of the diversity. Perhaps the best metaphor is running a magnet along the nails. How can all of the gems within the program be best brought to light? Even at the level of the project support, it will be necessary to strongly support and steer them to ensure relevance to the broader WEE sector.

What is the strongest policy narratives related to inclusive growth? Keeping in mind again, how evidence can be used, it will be important what GrOW's evidence is encouraging in terms of growth, macro-economic policies, other complementary mechanisms. At the moment, there is mixed and inconclusive evidence about the relationship between WEE and economic growth in the broader field. GrOW can lead the dialogue around what it takes to not only support inclusive growth but to conduct global research in a way that advances the same outcomes.

Key Recommendations

4.7 Recommendations to Grow Executive Committee

The main recommendations to the Executive Committee relate to GrOW program policy strategy and standard setting. There is also a recommendation for future research programs.

Recommendation 1: Segment policy narratives and stakeholders, notably,:

i. Strong policy narrative at the global level. e.g. factors related to occupational and sectoral segregation [global multi-laterals; academe];

ii. Evidence that provides a consistent narrative across contexts or regionally. e.g. care economy; transition pathways from education to work or informal to formal employment [regional and national bodies]

iii. Evidence that is more context specific and likely to be most relevant in-country. E.g. Skills training in Pakistan [ministries, private sector, NGOs]; and

iv. Evidence that is not strong enough quality to be publicized or to influence policy.

This segmentation of the policy narratives allows the Program as a whole to have differentiated strategies for project support, as well as policy engagement to continue to draw on relevant policy networks.

Recommendation 2: Do less of higher quality and differentiate outputs

It is highly recommended that both the quantity of outputs and research output guidelines be revisited and potentially reframed to reflect the nature, quality of the evidence and the segmented stakeholders. Otherwise, the program may be at risk of capturing a wide range of mixed quality outputs that are not well-positioned to influence. Reconsider, in particular, the emphasis on synthesis papers at the end of the program in favour of strong support to projects to positon and align toward common policy narratives.

With respect to standard setting, the standards that have been agreed upon are an important foundation. Any research published by GrOW should be defendable by common research principles of rigour (DFID, 2014) and uphold the balance of outcomes in the quality framework that has not only been important, but is a value-added for GrOW (Ofir et al., 2016). Given the diversity of projects in both nature and quality, it is advisable to consider differentiating them in terms of expectations of research outputs.

By the end of five years, the GrOW program expects the following the results:

- 60 empirical working papers on research themes 1 and 3, including on new methodology
- 8-10 Evidence Syntheses on research theme 2
- 25 publications in peer-reviewed journals and books
- 50 researchers well equipped for high-quality, policy relevant research
- 75 researchers achieve strengthened engagement with policy makers
- 20 peer-reviewed articles co-authored or authored by southern researchers
- 15 effective research uptake strategies at project level

- participation of 50 international policy makers in the research programme
- 25 synthesis papers, policy briefs, etc. to consolidate programme knowledge.

(Chapman, 2013, p. 15-16).

Overall, the GrOW program and Executive Committee will want to seriously consider whether the program should try to meet the quantity of deliverables expected. It is recommended that the Executive Committee seriously reconsider the level of outputs expected in favour of fewer focused high-quality outputs. As earlier shown, there is risk that projects are focused on "getting deliverables out" over ensuring coherence and quality. The program would greatly benefit from taking a step back and considering fewer, high quality research outputs that should likely vary by research project.

Since the projects are currently producing their early outputs, it may be helpful to have standards or guidelines that match the type of research they are doing. For example, country-level studies could have an integrated output that combines the qualitative and quantitative analysis with quality and a particular audience in mind. Academic research projects may focus more on peer-reviewed articles than policy briefs. In other words, every type of research output need not be expected from each research project Rather, consider different types of outputs against the segmentation of evidence and stakeholders for each.

For example:

i. *Peer-reviewed journal articles*. A peer-reviewed journal article is usually much more academic or theoretical in nature and highly focused. These should also combine, where relevant the qualitative and quantitative analysis and sound contextualization. The audience should be clear and relate well to the type of evidence. These might be more academic or meta-level scientific data across countries and be relevant in macro-economic policy circles.

ii. *Policy brief*. These can follow the guidelines already established which are helpful.

iii. *Briefing paper.* These can follow the guidelines already established as well with a caution that a great deal of variety may appear under this category. The program might consider that a "Briefing paper" be geared to practitioners or private sector providers, for example.

iv. *Grey literature or draft paper*. There should be an output that reflects the quality of the output or perhaps the stage in the process. It may reflect a stage in the process or learning that did not reach a level of quality to be relevant to a particular audience. Perhaps these can be utilized in some form over time. There should be an output that helps to clearly name this situation.

The risk with working papers and synthesis papers is that there will be many of these without clear audiences and too long, too full of data and internally-focused to be useful. Also, that the treatment of them will be extremely varied. The risk for the program is a number of publications of highly mixed quality and use. If the outputs are well-focused, they will be of stronger impact at the program level too. For example, even a handful of briefing papers geared toward private sector providers of skill training and child care services may be more valuable and relevant widely than many working papers on the same topics.

One way to support greater focus is to request fewer outputs of a higher quality. At a minimum, each project should be able to produce at least one peer-reviewed journal article and 2-3 policy briefs or briefing papers targeted to strategic audiences. Of course, the outputs should be well-matched to the type of evidence and audience. So, the academic consortium may have more peer-reviewed journal articles that are appropriate to their audiences. Country-level research may have more briefing papers and perhaps only one peer-reviewed article. A project that may not be able to meet quality standards may only have grey literature and a briefing paper or note.

Recommendation 3: In future call for proposals, consider how design can enhance coherence

If possible, in future call for proposals, design for coherence. So, if the first call were to focus on systematic reviews, for example, future calls can be focused on analyzing the contextual nuances of these trends and gaps with a similar mixed-methods emphasis.

A second recommendation to support quality in research projects is to request as a deliverable the methodology and instruments that must be approved prior to field work and the next tranche of funding. This intervention point is perhaps the most critical in catching quality issues early enough to address them with restructuring and redesign, additional external supports if necessary. At this point, emerging team structure, capacity issues can be identified and addressed.

4.8 Recommendation to GrOW Program Management and Project Officers

For the GrOW Program Management and Project Officers there are a few pointed recommendations that are reinforced by earlier findings:

Recommendation 4: Continue to share what is working well

As the team has been doing, share both the outputs and the mechanisms of strong projects to support peer mentoring and providing projects that are having challenges with solid examples and strategies. It will be helpful to reinforce this once you have decided on your outputs.

Recommendations 5: Focus on gaps in the WEE sector and program-level policy narratives going forward.

There are clear policy narratives emerging. Work to form a coherent and quality body of evidence at the program level. This will mean lots of iteration with projects toward these aims. Just as strong projects engaged continuously with policymakers the program could be working with projects to harmonize these policy threads and narratives so that they speak as strongly as they can to the gaps identified in the sector. Keep an eye to the overall sector and help the project narratives emerging to align more effectively.

Recommendation 6: Support the projects that are having difficulties to focus much more and align with policy narratives.

For projects that require corrective measures, many issues had to do with coherence. Insist on coherence between qualitative and quantitative analyses even if that might mean fewer research outputs. Again, better to have fewer focused outputs that can really be used effectively. Ensure that project direction has very clear connections to WEE, to the program policy narratives and leverage points not only barriers.

This stage of capacity building may look a little different as projects may require different supports. Some projects will require more. Particularly in these situations, aim to support the in-country researchers and the local institutions to build their internal capacity to do solid policy-focused research. For those that may not be able to produce “quality” outputs by the end, this may be a continued relationship with the local research institution. What kinds of supports, skills best position them to build capacity for this program and going forward?

Recommendation 7: Balance a differentiation of policy influence and capacity needs with standards and accountability.

The movement toward consensus around standards is critical. Going forward, accountability to the standards will be important to ensuring the legitimacy and influence of the program overall. However, standards and accountability do not mean that the same expectations or supports need to be used for every project. Differentiation may be helpful both in terms of policy influence and capacity supports. In some cases, expectations for what is expected may warrant revisiting to allow different types of research and research partnerships to influence in the way best suited to their project and capacities.

In the same way, for some projects that require support, focus more on capacity building and the longer-term relationship and outputs rather than pushing immediate published outputs of poor quality. It is important to note that quality issues exist for both North and South based institutions and researchers from the Global North and South.

For those projects that require corrective mechanisms, continue with the frequent communication but document the process more frequently through PMRs. It is perfectly acceptable to demand outcomes, and a minimum of quality outputs given the public funds that have been provided for that purpose. It may be necessary to use stronger mechanisms and direct the researchers more to ensure these. Otherwise, the risk is that program overall lacks coherence and legitimacy.

Recommendation 8: Focus on outcomes for policy uptake and capacity building.

As earlier stated, it was often difficult to assess the capacity building and to a lesser extent, the policy uptake strategies, in part, due to the nature of the monitoring that focused more on outputs than outcomes and differentiated strategies and priorities. Reconsider the reporting but mainly dialogue on outcomes going forward. Support the projects to prioritize their policy and practitioner targets. These will likely be few, much fewer than the many stakeholders listed in their strategies. Those stakeholders may be useful at the communication and dissemination stage later.

Support projects to prioritize influencers and match their outputs based on gendered context analysis. How does change and influence work in that context and which stakeholders are best positioned to influence? It may not be the Ministry or the Ministry most closely linked to the issue. It may be organizations associated with the women’s movement in that country. Model with the projects that have done this process effectively as many have. Continue to support peer mentoring in these regards.

5 Lessons Learned

5.1 About research quality and evidence

The varied, sometimes conflicting, views about research quality and evidence are helpful to explore since they represent tensions in emphasis in the broader field. This program throws these debates into relief.

Where one perspective is that rigour is the pre-condition for policy uptake and capacity building there is another view that is not exclusive but overlapping. That is, there are tensions between the three outcomes and therefore seen more as a balancing act than a trajectory. A prioritization of rigour risks reliance on established institutions and researchers, therefore, doing little to change existing power structures.

What is the role of research in redressing power dynamics? This was one of the questions raised during dialogue within the Mid-Term Evaluation. An external stakeholder warned that the current move more broadly toward big data and experimental research in the name of rigour is making these nuanced issues of WEE more and more the area for established economists and universities. At the same time, producing poor quality work does little to advance either the positive outcomes of WEE or the capacities. There is a balance to be sought and GrOW has managed well to do that, for the most part, with a remarkably diverse mix of disciplines, types of research, types of research institutions.

These views are described as overlapping in that most agree that there are tensions between the outcomes and that they can be mutually reinforcing. While a minimum of rigour is necessary to have policy uptake and capacity building, it cannot be assumed that rigour leads to policy uptake and capacity building without an intentional approach, reflection and adaptation.

As many, including in-country researchers emphasized, building capacity well takes time, intentional research process design. It is critical, given the power dynamics and gender norms at play in this area, that the questions and researchers continue to be diverse in all of the backgrounds that intersect with gender. Some researchers also pointed out that this is a long-term process. Economists from the North, as part of the portfolio, certainly have a role to play particularly since that audience is critical for future change in power structures. It may be questionable if the portfolio leans too heavily to this or any particular segment. A program such as this, in its diversity, is an important step in building the relationships and providing researchers and institutions with opportunities to expand the analysis and the dialogue so critical for change.

5.2 About monitoring and evaluation

The outcome mapping used in the evaluation and adapted from IDRC's framework, is part of a larger systems approach to monitoring and evaluation that prioritizes dialogue, process and not only results but perceptions and motivations. Particularly for formative Mid-Term Evaluations such as this one, part of the role of the MTE is to stimulate important dialogue around key issues, tensions and trade-offs such as the one in the previous section.

Outcome mapping facilitates processes and discussions by rendering a pathway transparent and allowing stakeholders to discuss it. Through this process, it is usually discovered that assumptions made about the pathway may not be correct. In this way, there is learning and adaptation about what is working. For example, the indicators within the project assessment were adapted over time following the Mid-term workshop, for example. In the early results workshop with the Program team and the EC, there was lively discussion about project ratings not only about the accuracy of where they were placed but if "expectations" under expect to see, like and love to see were reasonable and realistic. In this way, through iteration, everyone develops a better understanding of what could and should be expected.

This process was even useful to dialogue around some of the dichotomies and challenge them. As one in-country researcher said, the South-based universities might be stronger. One of the "in-country researchers" was a Harvard history professor from Ghana. Someone challenged, how is he categorized? One researcher from a Northern university admitted to knowing nothing about what it meant to contextualize the research. Another highly established academic admitted to being "totally green" when it came to understanding how to do policy uptake. Categories and segmentation can be helpful as this report illustrates. However, it is helpful to be reminded of nuance, so that we don't dangerously collapse ideas such as in-country researchers as lacking capacity. Instead, we try as we must in the WEE sector, to capture complexity and nuance as best we can.

6 Conclusions

The GrOW Program is positioned to add value and address gaps in the broader field of Women's economic Empowerment. To do so requires strong and strategic focus on the overall GrOW Program including how the projects are supported. It may mean directing projects more to ensure relevance to the sector, and that some projects are highlighted differently than others. Some projects where there are concerns about the validity and reliability of the research, may require a different approach with a focus on building individual and institutional capacity. Overall, it is critical that each project is focused on where they can best contribute to the larger body of evidence, policy narratives and dialogue. The metaphor going forward is a magnet over nails.

Appendix I References

Bandiera, O. and Natraj, A. (2013). “Does gender inequality hinder development and economic growth? Evidence and policy implications” in The World Bank Research Observer, 28(1), 2-21.

Buvinić, Mayra. and Furst-Nichols, Rebecca. (2014). “Promoting women's economic empowerment: what works?” in The World Bank Research Observer, lku013.

Buvinic, Mayra, Rebecca Furst-Nicols and Emily Courey Pryor. (2016). A Road Map for Promoting Women’s Economic Empowerment. United Nations Foundation. Exxon Mobil. New York, United States.

http://www.womeneconroadmap.org/roadmap

DFID. (2014) Assessing the Strength of Evidence, London, United Kingdom.

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf

Doepke, M. and Tertilt, M. (2014). “Does female empowerment promote economic development?” in National Bureau of Economic Research, No. w19888.

Duflo, E. (2011). Women's empowerment and economic development. National Bureau of Economic Research, No. w17702.

Elborgh-Woytek, M. K., Newiak, M. M., Kochhar, M. K., Fabrizio, M. S., Kpodar, K., Wingender, M. P. and Schwartz, M. G. (2013). Women, work, and the economy: Macroeconomic gains from gender equity. International Monetary Fund.

Eyben, R. et al. (2008). Conceptualizing empowerment and the implications for pro-poor growth, Institute of Development Studies at the University of Sussex, Brighton.

Global economic forum (2015). The Global Gender Gap Report 2015. World Economic Forum. Geneva, Switzerland.

Growth and Economic Opportunities for Women. (2013). Literature Review to Inform the DFID-IDRC-Hewlett Foundation Research Program on Women’s Economic Empowerment, Gender Equality and Growth in Low Income Countries. International Development Research Centre. Ottawa Canada

International Labor Organization. (2016). Women at Work: Trends 2016. Geneva, Switzerland.

Kabeer, N (2015). Gender, Poverty and Inequality: A brief history of feminist contributions in the field of international development. Gender and Development 23:2, pp. 189-205.

Molyneux, M. (1985). Mobilization without Emancipation: Women’s interests, the state and revolution in Nicaragua. Feminist Studies 11(12), Summer, pp. 227-254.

Ofir, Zenda, Thomas Schwandt, Colleen Duggan, Robert McLean. (2016). Research Quality Plus: A Holistic Approach To Evaluating Research. International Development Research Center. Ottawa, Canada.

Popay, Jennie, Helen Roberts, Amanda Sowden, Mark Petticrew, Lisa Arai, Mark Rodgers, Nicky Britten (2006). Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. for guidance on combining evidence from quantitative and qualitative sources in reviews.

Taylor, G. and Pereznieto, P. (2014). Review of evaluation approaches and methods used by interventions on women and girls' economic empowerment. Overseas Development Institute, (ODI), London.

Waddington , Hugh, Howard White , Birte Snilstveit , Jorge Garcia Hombrados , Martina Vojtkova, Philip Davies , Ami Bhavsar , John Eyers , Tracey Perez Koehlmoos , Mark Petticrew , Jeffrey C. Valentine & Peter Tugwell (2012) How to do a good systematic review of effects in international development: a Tool kit, Journal of Development Effectiveness, 4:3, 359-387, DOI: 10.1080/19439342.2012.711765 . To link to this article: http://dx.doi.org/10.1080/19439342.2012.711765 accessed August 29, 2016

Appendix II Evaluation Team Biographies

Nanci Lee: Team Leader

Ms. Nanci Lee (MSc) has accumulated over 20 years of experience working in microfinance with an emphasis on gender, women's economic empowerment, market systems analysis, market research, demand-driven products, double-bottom line performance monitoring and evaluation. Currently she is conducting a systematic review on financial education interventions aimed at HIV-vulnerable youth, orphans and vulnerable children. She worked as Program Coordinator for VLS Associates with funding from Bill and Melinda Gates Foundation, where she was responsible for the design of SAVIX database for monitoring (mostly informal) savings groups worldwide. Prior to this, she led and coordinated Coady Institute's Community-Based Microfinance for five years and continues to act as an Associate Instructor and Researcher. In 2007-2009, Ms. Lee has overseen two large multi-country research studies. The first funded by Ford Foundation, used case-studies of cooperatives, financial associations and savings groups in Africa, Asia and Latin America. Key thematic areas included member governance, networking and linkages, and regulation and supervision. The second multi-country study involved the implementation of a market research and social performance system in Mali, Algeria and Namibia including links between informal coping strategies and asset building and access to formal finance. All of her work has focused on savings products and the link between savings and asset building in households and members, particularly women. Ms. Lee also has a strong expertise in project evaluation. She has reviewed and evaluated microfinance and cooperative programs in Angola, Ethiopia, Tanzania, Malawi, Kenya, Cambodia and Canadian credit unions. She worked on policy and regulation for member-owned microfinance with the Bank of Angola and with an Angolan commercial bank in Angola on linkages with producer associations.

Rima Slaibi: Consultant

Ms. Rima Slaibi (M.A) has a solid and diversified experience in conducting evaluations. As part of her involvements on projects, she conducted and participated to evaluations in a variety of fields such as employability and entrepreneurship (Oxfam Quebec, Drosos/EFET), training and capacity development (World Food Program, Oxfam, SOCODEVI, Laval University), and gender (UN-Women, Oxfam). On employability and entrepreneurship, Mrs. Slaibi acted as lead evaluation expert for the evaluation of EFET's (Education for Employment – Tunisia) "Moustakbali" project. She also acted as an Evaluation Advisor for the Elaboration of an Evaluation Plan for the Oxfam Québec Youth Internships Program. On Research, her areas of expertise in evaluation have been informed by her research and publications. On Gender, Mrs. Slaibi acted as Lead Evaluation Advisor for conduct of a Baseline study for a Gender analysis and Environmental analysis for the Oxfam "Community-Led Integrated Water Resource Management project in Vietnam". She also acted as Lead Evaluation Advisor for the conduct of an Impact Evaluation of a Gender equality project of Oxfam in Vietnam entitled "Gender Talks-Women's Leadership and Economic Empowerment". Mrs. Slaibi also participated in the conduct of two evaluations of UN-women's contribution to United Nations system coordination on gender equality and women's empowerment, at global and ECA regional level.

Halcyon Louis: Consultant

Ms Halcyon Louis (MSc) is an international development consultant with over 14 years of experience in progressive socio-economic research, including project monitoring and evaluation, organisational assessment, and policy development. Her experience has been acquired in 22 countries across Africa, Asia, the Caribbean, North America and Europe. She is a development economist by training, and manages a work portfolio centred on a social development thematic, with emphasis on sub-themes that include youth (asset development); poverty reduction; education; and gender dynamics. She has extensive experience in the application of popular participatory methods to social research design, and multi-partner project evaluations, and is versed in the application of qualitative and quantitative analytical techniques, and in technical reporting. Ms Louis has worked collaboratively on programme design and implementation as an active member of project steering groups and advisory committees. She has functioned as the programme manager and team leader of contracted consultancy assignments, as well as the lead coordinator of capacity-building and information dissemination workshops. Ms Louis has also supported development agencies, and organisations in the public, private and non-governmental sectors, to enhance their development effectiveness through results-based management. She has also conducted evaluations that were designed to assess program/ project impact. Included among Ms Louis' diverse clientele are: multilateral organisations and development banks; government and inter-government agencies; donor networks; non-governmental organisations and non-profit agencies; and charitable foundations and grant-makers.

Appendix III Documents Reviewed

GrOW projects	Research Outputs and Early Evidence Reviewed	Research uptake strategy	Technical reports	Budget	Project Monitoring Reports
107807 Women's Early Labour Market Transitions in Sub-Saharan Africa	Educational attainments, childbirth outcomes among young women in six selected Sub-Saharan African countries: An analysis based on the DHS	Y	3 tech reports 2015 and 2016 (2)	777,600	3 PMRs (July 2015, January 2015, April 2016)
107808 Improving Childcare Options to Create Better Economic Opportunities for Women in Nairobi Slums	Methodology Instruments 2 briefing notes	Y	Inception report (in lieu of a tech report for 2015) and 2 tech report for 2016	952,466	2 PMR (January 2015 and April 2016)
107809 Understanding the Effect of Skills Training on Women's Economic Opportunities in Pakistan	Methodology Preliminary findings	Y	Inception report + 1 technical report in August 2015 + technical report from 2016	904,600	-
107811 Assessing the Impact of Cash Transfer Programs on Women's Empowerment in Tanzania	Assessing the impact of cash transfer programs on WE in Tanzania Pre-Analysis Plan A qualitative assessment of WE in a CCT programin Tanzania Qualitative findings on the /five domanis of empowerment Quantitative report	Y	3 tech reports 2015 and 2016 (2)	816,300	-
107818 Together We Can: Assessing the Impact of Women's Action Groups on Social Change in India	Methodology Instruments	Y	3 tech reports 2015 and 2016 (2).	902,300	3 PMRs
107819 The Influence of Affordable Daycare on	The effect of an Affordable Daycare Program on Health and Economic Well being in	Y	3 tech reports 2015 and 2016 (2)	1,000,171	2 PMRs

GrOW projects	Research Outputs and Early Evidence Reviewed	Research uptake strategy	Technical reports	Budget	Project Monitoring Reports
Women's Empowerment in India	India: A Cluster-Randomized impact Evaluation Study (Baseline Report) A pracial Guide to Measuring Women's Empowerment Research Protocol Childcare and women's health, social and economc outcomes in low and middle-income countries: A systematic Review Unpaid work is associated with mental distress: evidence from women in rural India				
107820 Understanding Women's Experiences in Artisanal and Small Scale Mining in Central and East Africa	Sexual Violence, Conflict Minerals and the "Economics of Appearance" Book chapter draft; working paper (2) Instruments Methodology	Y	3 tech reports 2015 and 2016 (2)	996,087	-
107821 Addressing the Barriers to Young Women's Economic Empowerment in Bangladesh	Draft qualitative journal article; cost-based analysis; working paper	Y	3 tech reports 2015 and 2016 (2)	970,700	-
107850 Identifying Post-War Growth and Economic Opportunities for Women in Sri Lanka	Literature review Instruments	Y	3 tech reports 2015 and 2016 (2)	719,000	2 trip reports + an email

GrOW projects	Research Outputs and Early Evidence Reviewed	Research uptake strategy	Technical reports	Budget	Project Monitoring Reports
107852 Balancing Unpaid Care Work and Paid Work in South Asia and sub-Saharan Africa	Methodology Instruments	Yes, with the calendar of activities	1 inception report, 1 tech report from 2015 and I tech from 2016	992,700	-
108111 Pathways for shared prosperity: Understanding the links between WEE and growth	Very useful website 15 papers Global - The impact of gender inequality in education on economic growth: protocol for a systematic review and met-analysis - Drivers of Gendered sectoral and occupational segregation in developing countries - A Flow Measure of Missing Women by Age and Disease Pakistan - The impact of fundamentalist terrorism on school enrolment: Evidnece form NW Pakistan India - Intra-household gender disparity in school choice: Evidence from Private Schooling in India - Intra household gender disparity in school choice: Evidence from private schooling in India Jordan	Y	3 tech reports 2015 and 2016 (2)	2,255,300	-

GrOW projects	Research Outputs and Early Evidence Reviewed	Research uptake strategy	Technical reports	Budget	Project Monitoring Reports
	- Does women's labour force participation reduce domestic violence? Evidence from Jordan Nepal - Apprenticeship training and female EE: Quasi experimental evidence from Nepal Colombia - Sorting through Affirmative Action: Three Affirmative Action experiments in Colombia South Africa - The impact of affirmative action on occupational segregation by gender in South Africa (full paper) (PPT preliminary findings) - The determinants of female informal sector employment - Gendered constraints to job search - The tenure and experience profiles of different demographic groups - The unintended consequences of education policies on SA participation and unemployment -Migration and poverty transitions				

GrOW projects	Research Outputs and Early Evidence Reviewed	Research uptake strategy	Technical reports	Budget	Project Monitoring Reports
	- Gender, fertility and health Spatial patterns of employment				
108112 Making Growth Work for Women in Low-income countries	Women's EE: A Review of Evidence on Enablers and Barriers National Income changes and the Empowerment of Women within the Household	Yes	1 inception report + 1 tech report from 2016	2,166,300	-
108114 Bringing to light the role of the extractive industry on women's economic opportunities in Côte d'Ivoire and Ghana (GrOW)	None provided. Review of email correspondence with PO.	Brief ppt	Inception report + Technical report of July 2016	336,800	PMR from April 2016
108115 The impact of women's political representation on economic growth and WEE in Africa	None provided.	Yes, brief	Inception report + tech report 2016	337,100 No outputs expected	-

Note: shaded colours indicate grouping by category. No shaded are the strong across outcomes. Lightly shaded can be improved with corrective mechanisms and the darkest shaded were weak across outcomes.

1. **GrOW program documents:**

- GrOW annual reports for 2014 and 2016 submitted to DFID
- DFID annual reviews for 2014, 2015 and 2016
- GrOW overview (initial and 2016)
- GrOW call documents for theme 1, 2 and 3
- GrOW project status (basic project data) as of August 2016

Appendix IV GrOW Detailed Evaluation Matrix

Individual Research Project Quality

	Evaluation Question	Indicators	Data Sources
1. RESEARCH QUALITY	*Focused on the what of the research itself*		
Rigour	1.1. What is the rigour of the research project based on early outputs? Note: Rigour can be assessed for projects for which no methodological risks identified in PAD and are at least reporting preliminary results or a draft report	1.1.1 Internal validity (minimization of risk and bias errors) mentioned or evident 1.1.2 Appropriate controls mentioned or evident 1.1.3. Reliability (authors have considered study limitations and alternative interpretations) mentioned or evident 1.1.4 Generalizability mentioned or evident 1.1.5. Clear link to outcomes mentioned or evident 1.1.6 Acceptance to a respected peer-reviewed journal	For each of the indicators the following sources will be used in combination: • Methodology and research instruments • Draft preliminary results, working papers, policy briefs (at varying stages of completion) • Tool 1: Document Review. Background documents/Technical reports • Tool 3 and 4: Principal Investigators and Lead Researchers (Survey; Interview)
Positioning for Rigour	1.2. What is the positioning for rigour of the research project based on early outputs? Note: For research projects identified in PAD as having some methodological risks **and/or** at earliest stages of completion	1.2.1 Coherent theory of change/hypothesis mentioned or evident 1.2.2 Appropriate and valid research methods and methodology including sampling mentioned or evident 1.2.3 Outcomes are identified (including how they will be measured) 1.2.4 Quality control mechanisms are defined and there is evidence of use 1.2.5 Evidence of improvements in methodological design 1.2.6 Use of a research protocol such as Campbell, Cochrane etc	For each of the indicators the following sources will be used in combination: • Methodology and research instruments • Draft presentations • Tool 1: Document Review. Background documents/Technical reports • Tool 3 and 4: Principal Investigators and Lead Researchers (Survey; Interview)
Legitimacy	1.3. What is the legitimacy of the research in terms of its reputation, grounding in local context and attention to gender and ethics?	1.3.1 Ethics standards and strategies in place 1.3.2 Evidence of gender analysis in design, process and research outputs 1.3.3 Strategies to ensure that research is grounded in local context and realities 1.3.4 Reputation of research leads and team	1.2. 1. Ethics proposal/agreement-universities. Disclosure statement to human subjects. Tool 3: Grow Pre-Meeting Survey of Research Projects 1.2.2 Tool 1. 1.2.3 Tool 3 and 4: Principal Investigators and Lead Researchers (Survey; Interview) 1.2.4 Body of evidence on which they are building. Tool 10: Key Sectoral Stakeholder Key Informant Interview
Relevance	1.4. To what extent does the research fill a gap in the sector with respect to innovation or importance?	1.4.1 Strategies for identifying gaps and value-added 1.4.2 Relevance of research to sector 1.4.3 Schematic map of GroW research within WEE sector	1.3.1.1 Tool 1. Tool 3 1.3.2.1 Tool 10 1.3.3.1 Systematic and literature reviews on WEE. Scan of the sector. Tool 9: Sectoral Evidence Mapping Framework
Synthesis	1.5 What is the indication of quality by this research project based on its current trajectory?	1.5.1 Presented in a diamond chart form with uptake and capacity building to contrast dimensions along a trajectory. Narrative write-up will highlight green, yellow and red lights.	

Effectiveness of Research Project Positioning

	Evaluation Question	Indicators	Data Sources
2. RESEARCH POSITIONING	*Focused on the individual research project strategies (the how)*		
Effective Project Management	2.1. How effectively is the project managing to ensure quality, risk management and adaptation to emergent learning?	2.2.1 Most successful elements of research 2.2.2 Areas of challenge, delay. Evidence of risk management 2.2.3 Evidence of where research process or product changed or needed to be adapted 2.2.4 On time (completed deliverable against intended) with "reasonable" delays 2.2.5 On budget (against intended) with "reasonable" variance 2.2.6 Strategies for ensuring coherence and comparability (where applicable)	Tool 1: Document Review Tool 2: Project Officers Tool 3: Principal Investigators and Lead Researchers (Survey) Tool 4: Principal Investigators and Lead Researchers (Interview) Follow up pointed interviews with individual researchers
Effective Positioning for Uptake	2.2 How effectively is the project positioning the research for uptake by influential decision-makers?	2.3.1 Timeliness (current; piggyback key events) 2.3.2 Accessibility in dissemination strategy (outreach platform/mechanism (online; free; etc.); language) 2.3.3 Appropriateness of uptake strategy in targeting influential decision-makers (actionable; strategic; audience-focused; specificity of uptake and influence) 2.3.4 Early evidence of uptake (citations; policy events or dialogue)	Tool 1 Tool 2 Tool 3 Tool 4 Follow up pointed interviews with individual researchers Tool 10: Key Sectoral Stakeholders (Interview)
Effective Strategy for building In-Country Capacity	2.3 How effectively is the project creating opportunities for building in-country capacity?	2.4.1 In-country researcher participation in research design 2.4.2 Appropriateness of capacity building strategy 2.4.3 In-country researcher participation in events 2.4.4 In-country researcher sharing knowledge at events 2.4.5 In-country researcher sharing knowledge on learning platforms 2.4.6 Authorship or writing contribution opportunities for local researchers 2.4.7 Extent of local authorship and co-authorship	Tool 1 Tool 2 Tool 3 Tool 4 Follow up pointed interviews with individual researchers
Synthesis	2.4 How well positioned is the research project to achieve quality evidence, effective research uptake and built capacity of in-country researchers given its current trajector?	1.5.1 Presented in a diamond chart form with quality dimensions to contrast dimensions along a trajectory. Narrative write-up will highlight green, yellow and red lights.	

Research Program Positioning

	Evaluation Question	Indicators	Data Sources
3. RESEARCH PROGRAM POSITIONING	*Focused on the extent to which overall Research Program is well-positioned to achieve its three key outcomes by end of Program (future-oriented)*		
Enhanced high quality evidence available for decision making (Outcome 1)	3.1 How effectively is the Research Program positioned to ensure a high quality body of evidence is left that adds value to the sector?	3.1.1 Synthesis of individual projects scores on quality and risk 3.1.2 Relevance of Research to sector and potential influence	3.1.1.1 Individual Project Assessments 3.1.2.1 Team Assessment 3.1.3.1 Tool 9 and 10: Evidence mapping and interviews with key sectoral stakeholders 3.1.4.1 Mid-Term Workshop (Germany)
Enhanced uptake of research by key influential decision-makers (Outcome 2)	3.2 How effectively is the Research Program positioned to ensure influential decision makers will make use of the research?	3.2.1 Synthesis of uptake strategies 3.2.2 Review of sectoral learning events, platforms, context in uptake strategies to ensure appropriate range of forums 3.2.3 Evidence of tailoring/differentiating and targeting key decision makers 3.2.4 Evidence of stakeholder mapping and early engagement with key decision makers	3.2.1.1 Individual Project Assessments 3.2.2.1 Team Assessment 3.2.3.1 Tool 9 and 10: Evidence mapping and interviews with key sectoral stakeholders 3.2.4.1 Mid-Term Workshop (Germany) 3.2.5.1 Tool 1: Document Review
Enhanced in-country capacity (Outcome 3)	3.3 How effectively is the Research Program positioned to have built in-country research capacity by the end of the Research Program?	3.3.1 Ratio and relation of local researchers to other researchers 3.3.2 Rate and depth of local participation relative to other researchers in learning platforms, events and publications 3.3.3 Evidence of built local research capacity	3.3.1.1. Tool 1: Document Review. List of publications and authors 3.3.2.1 Tool 1: List of events, panel speakers, presenters, participants 3.3.3.1 Follow up interviews with key in-country researchers 3.3.4.1 Mid-Term Workshop (Germany)
Research Program Management	3.4 How effectively has the Research Program managed for quality, risk and adaptation?	3.4.1 Synthesis of individual project scores and quality, risk, management 3.4.2 Examples of strongest results, practices and strategies across Research Program 3.4.3 Analysis of changes made to Research Program and adaptation to risks, changes 3.4.4 Setting of common terms understanding - success 3.4.5 Frequency and quality of support to research projects 3.4.6 Research project perception of value added of this funding and structure	3.4.3.1 DFID Program Reviews 3.4.3.2 Mid-Term Workshop (Germany) 3.4.3.3 Tool 3 and 4 (Interviews with Principal investigators and lead researchers) 3.4.3.4 Tool 8 (Interview with Executive Committee)
Synthesis	3.5 How well is the overall Research Program positioned to achieve its three outcomes by the end of the program?	Presented in a schematic diamond form form where individual research projects can be plotted with small dots, outcomes as large dots to give an overall graphic.	

Appendix V Summary of Meetings and Interviews

Respondent(s) and Position(s)	Date	Interviewers	Modality
Project Officers/ Project Managers			
Alejandra Vargas Garcia, Lead on Research Uptake Strategies and Communications	Sept. 19 – 22, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview plus follow up
Arjan de Haan, Team Leader, Employment and Growth	Sept. 19 – 22, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview
Edgar Rodriguez, Project Officer	Sept. 19 – 22, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview plus follow up
Flaubert Mbiekop, Project Officer	Sept. 19 – 22, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview plus follow up
Madiha Ahmed, Project Officer / Overall Program Coordinator	Sept. 19 – 22, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview plus follow up
Martha Melesse, Senior Project Officer	Sept. 19 – Dec 16, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview plus follow up –
Paul Okwi, Project Officer	Sept. 19 – 22, 2016	Nanci L; Rima S; Halcyon L	– Telephone interview plus follow up
Robert McLean, Senior Program Specialist - Evaluation GrOW Program Team	Dec 16, 2016	Nanci L; Halcyon L	– Preliminary Results Presentation, IDRC
Principal Investigators			
All	October 10-October 31, 2016	Nanci L; Rima S	- Survey by email
Andy McKay Jane Mariara	Oct.17-18, 2016	Nanci L Nanci L	– IDRC Conference Meeting

Respondent(s) and Position(s)	Date	Interviewers	Modality
107807			– Follow up Survey/ Email communication
Arijit Nandi Parul Agarwal 107819	Oct. 16, 2016 Oct. 17-18, 2016	Nanci L; Rima S Rima S Nanci L	– Phone interview – IDRC Conference Meeting – Follow-up
Caroline Kabiru Frank Grimard 107808	Oct. 18, 2016	Nanci L Nanci L	– IDRC Conference Meeting – Follow-up
Deepta Chopra Shraddah Chigateri Jenipher Torch 107852	Oct. 17, 2016	Nanci L Nanci L Nanci L	– Telephone interview – IDRC Conference Meeting – Follow-up
Doris Buss Gisèle Eva Côté Abby Sebina-Zziwa 107820	Oct. 13, 2016 Oct. 17, 2016	Nanci L Nanci L Nanci L	– Telephone interview – IDRC Conference Meeting – Follow-up
Elizabeth Peters Nan Astone 108112	Oct. 18, 2016	Rima S Nanci L	– IDRC Conference Meeting – Follow-up
Flora Myamba Abel Kinyondo Samuel Wangwe	Oct. 17, 2016	Rima S Halcyon L	– IDRC Conference Meeting – Follow-up
Janneke Pieters Stephan Klasen Servaas van der Berg Marcela Ibanez 108111	Oct. 18, 2016	Nanci L	– IDRC Conference Meeting – Follow-up
Sophia du Plessis Sahar Parsa	Oct 17, 2016	Nanci L	– IDRC Conference Meeting

Respondent(s) and Position(s)	Date	Interviewers	Modality
108115			– Survey/email
Jyotsna Jha Skreekanth Mahendiran 107818	Oct 13 – 18, 2016	Nanci L Nanci L; Rima S; Nanci L; Halcyon L	– Telephone interview – IDRC Conference Meeting – Follow-up
Ramani Gunatilaka Ranmini Vithanagama Kethaki Kandanearachchi 107850	Oct. 18, 2016	Rima S Nanci L; Halcyon L	– IDRC Conference Meeting – Follow-up
Nan Astone Elizabeth Peters 108112	Oct 18, 2016	Rima S Nanci L	– IDRC Conference Meeting – Follow-up
Rachel Glennerster Shahana Nazneen	Oct. 17, 2016	Rima S; Nanci L; Halcyon L	– IDRC Conference Meeting – Follow-up
William Baah-Boateng Silvère Konan	Oct. 18, 2016	Rima S; Halcyon L	– IDRC Conference Meeting – Follow up
Executive Committee			
Sue Szabo, IDRC	Oct. 19, 2016	Nanci L; Rima S	Mid-Term Workshop
Helena Choi, Hewlett Foundation	Nov. 7, 2016	Nanci L; Halcyon L	Telephone interview
Lina Chapman, DFID	Nov. 10, 2016	Nanci L	Telephone Interview
Madiha Ahmed, Program Manager	Nov. 23, 2016	Nanci L	Telephone Interview
In-country Researchers			
Stella Muthuri	Dec. 6, 2016	Nanci L	– Telephone interview
Raymond Atuguba	Jan 16, 2016	Nanci L	– Telephone interview

Respondent(s) and Position(s)	Date	Interviewers	Modality
Ammar Malik, LEAD	Dec 16, 2016	Nanci L; Halcyon L	– Telephone interview
Binate Namizata, CIRES	Dec 17, 2016	Nanci L; Halcyon L	– Telephone interview
Claude Wetta	Jan. 11, 2016	Nanci L; Halcyon L	– Survey/ Email communication
Danesh Jayatilaka	Dec 18, 2016	Nanci L; Halcyon L	– Telephone interview
Kethaki Kandanearachchi	Dec. 15, 2016	Nanci L; Halcyon L	– Telephone interview
Olivia Kwapong	Jan 10, 2017	Halcyon L	– Survey/ Email communication
Shreekanth Mahendiran	Nov. 26, 2016	Nanci L; Halcyon L	– Telephone interview
Stella Muthuri	Dec. 23, 2016	Nanci L	– Survey/Email communication
Manuel Silva Santos	Dec. 28, 2016	Nanci L	– Telephone Interview
Fred Kisekka	Jan. 10, 2017	Nanci L	– Phone interview
Anoushaka Chandrashekar	Jan. 11, 2017	Nanci L	– Phone interview
Emmanuel Akyeamppong	Jan. 18, 2017	Nanci L; Halcyon L	– Survey; email correspondence
Oyindamola Popoola	Dec. 14, 2016	Nanci L	– Phone interview
Adalbertus Kamanzi	Jan. 10, 2017	Nanci L	– Phone interview
External Stakeholders			
Jennefer Sebstad (independent, formerly USAID)	Jan. 2, 2017	Nanci L	– Phone interview
Caren Grown (World Bank)	Jan. 4, 2017	Nanci L	– Phone interview
Elizabeth Vazquez (WEE Connect)	Jan. 4, 2017	Nanci L	– Phone interview
Linda Jones (MEDA, SEEP Network)	Jan. 6, 2017	Nanci L	– Phone interview

Respondent(s) and Position(s)	Date	Interviewers	Modality
Karen Stefiscyn (Africa Energy Projects, formerly Centre for Human Rights, University of Pretoria)	Jan. 5, 2017	Nanci L	– Phone interview
Abigail Hunt (Overseas Development Institute);	Jan. 10, 2017	Nanci L	– Phone interview
Naila Kabeer (London School of Economics)	Jan. 13, 2017	Nanci L	– Phone interview
Sophia Mwakagenda (Member of Parliament, Tanzania, formerly gender activist)	Jan. 16, 2017	Nanci L	– Survey completed and emailed
Krista Jacobs (Gender Advisor, USAID)	Feb. 6, 2017	Nanci L	– Phone interview
Srilatha Batliwala	Feb. 27, 2017	Nanci L	– Phone interview

Appendix VI Tools and Interview Protocols

Tool 1: GrOW Research Program Document Review

Introduction

The objective of the document review is to provide a complete profile (and working typology) of the research projects based on level of completion and methodological risks earlier identified. The document review also includes all elements of the assessment tools and where to find them in specific documents. This provides a basis for capturing key content in the reporting documents (Proposals; PAD, Technical Reports; PMR, Uptake Strategies) related to key evaluation criterion. In concert with the project manager interviews and survey results, this review will help to guide specific questions of clarification and nuance for more in-depth interviews with principal investigators.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identify lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

[See Excel Attachment]

Tool 2: GrOW Research Project Officer Interview Guide

Introduction

This specific interview will, in combination with the research project survey information, provide a detailed profile of each research project. This interview will also provide the evaluation team with some guidance about how to access critical information in the different forums, particularly more nuanced information related to project adaptation, learning, capacity building, governance and team dynamics. Both the survey and interview results will determine the focus of more in-depth key-informant interviews in Berlin with research project investigators.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identify lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

Interviewee

Full Name: Gender:

Position: Responsibility/GROW:

General

1) Briefly describe your role and key areas of responsibilities related to the GROW Research Program.
2) What are the most promising or exciting elements of the research project so far in terms of advancing the sector? You might speak to quality /importance of evidence, research methodological approach, building capacity building of in-country researchers, policy uptake.
3) Briefly describe the governance structure of the research project that you oversee? (Include: team composition; international/in-country researchers, distribution of responsibilities, etc.)
4) In what ways has this structure enhanced or challenged the building of in-country research capacity either formally or informally?
5) What are some of the other drivers of building capacity of in-country researchers from your perspective?
6) What have been some of the key challenges related to building capacity of in-country researchers?
7) How would you advise the evaluation team to get to potentially sensitive issues of team governance, local capacity (as understood by the local researchers), power dynamics?

8) Is there someone, in particular, within the research project team that you recommend that we speak with to have a complete understanding of the process and unfolding outcomes?

9) Is it possible to give us a broad sense of whether this research project is on time and on budget?

10) Are there delays in the research project that you would consider unreasonable or likely to affect ultimate quality? If so, can you elaborate? Can you talk about how you've addressed these issues with the project?

11) In the PAD, the following risk were identified with the research project (e.g. spending more time on knowledge generation than policy uptake; not being able to effectively build and sustain team energy). Can you comment on your perspectives of the project's ability to manage these risks?

12) What are some elements of quality for research particularly for non-experimental research that is not peer-reviewed?

13) What have been some of the key challenges related to positioning for quality evidence?

14) What have been some key challenges related to positioning for effective uptake by key influencers?

15) What are some of the drivers of research uptake by key influencers from your perspective? To what extent has the research project taken these into consideration?

16) Can you comment on areas where the project has shown signs of learning and adaptation?

17) Anything else that would be helpful for the evaluation team to understand about the research project?

Adaptation (Uptake Strategies Lead):

1) Can you give examples of a few of the stronger projects in terms of uptake strategies? What makes them strong?

2) Can you give examples of a few of the weaker projects in terms of uptake strategies? What makes them weak?

3) Can you describe differences between the various governance structures in terms of uptake strategies or risks around them? (North-led with South-based partners; South-led country; South-South)

4) What are the drivers of effective research uptake that you are noticing?

5) What have been some of the key challenges that you have seen with research uptake (understanding that these are processes underway)?

6) What are some of the success measures for effective uptake by key influencers?

7) What would happen if a project largely failed to deliver on their uptake strategy?

8) Is there a set % allocated to each project for uptake?

9) Is there a GROW mandate to encourage either wide uptake and outreach or focused targeting? On what does the level of breadth and depth in uptake strategy depend?

10) In the DFID Annual Review of 2016, the challenge of having common understandings of key research outputs was noted. To what extent do you share this observation? Why or why not?

Tool 3: GrOW Pre-Meeting Survey of Research Projects

Introduction

> The purpose of the survey is to complement and clarify research project strategies already reported in project reports. The aim is not to repeat information already provided but to clarify project priorities as well as have information in a way that allows consistent (not comparable) treatment across highly diverse projects. This information will be used as a basis for interviews in the Mid-Term Workshop.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to commissioned to conduct a mid-term formative evaluation of GrOW. In the original Terms of Reference, the evaluation was expected to focus on (1) mechanisms in place for assuring research quality (2) strategies in place for enhancing research accessibility and for positioning research for uptake by scholarly, practice and policy communities; and (3) application of ethical research practice and integration of gender analysis.

The Universalia Management Group, founded in Montreal Canada in 1980, is a management consulting firm and recognized leader in monitoring and evaluation including multi-country and multi-stakeholder research. This team brings a solid background in both practice and research in women's economic empowerment and will compliment their experience with perspectives of key stakeholders in the sector.

Given how early most of the projects are in producing research outputs, the evaluation is formative, focused on assessing emergent findings. The aim is to use these findings to inform the project and program's positioning to achieve sectoral influence. The evaluation also aims to stimulate constructive dialogue about what makes quality research for women's economic empowerment.

The evaluation will review all of the research projects to assess how well they are positioned for quality (rigor, relevance, legitimacy). These quality aspects are combined with an analysis of project positioning for uptake and in-country research capacity. The evaluation will also provide considerations related to positioning of the overall research program in terms of the current program trajectory to achieve its key outcomes.

Please complete and send the survey by **October 14, 2016**. Any specific questions about this evaluation process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

Interviewee

Full Name:

Position: Institution (Country):

I. General

1. Please confirm your current state of completion on the overall research project as at September 30, 2016 by marking a large X in the stage that best reflects your current status:

Overall stage of completion	Methodology and instrument design complete	Baseline data collection complete	Preliminary results or beta testing on datasets (not yet publicly shared)	Second and third rounds of data collection in process (if applicable)	First draft paper, article or brief (but none yet publicly shared)	At least one working paper, article or brief finalized (publicly shared)

II. Positioning for Quality

2. How did you identify the importance (relevance; innovation) of your current research areas in light of gaps in the overall field of women's economic empowerment?

- Systematic review
- Literature review (informal but not a publishable output for GROW)
- Literature review (formal output for GROW)
- Building on and identification through existing knowledge and body of work by lead researchers on the team
- Other (please specify):

3. What strategies are you using to ensure rigor and quality?

- Experimental research design
- Independent review of study protocol
- Peer reviewed papers/articles
- Quality control protocol such as Campbell Collaboration
- Other (please specify):

3. a) Have you completed an ethics protocol through your research institution or university? YES/NO

b) If not, how are you ensuring ethical standards are in place?

4. Gender analysis is, of course, central to any research on women's economic empowerment. Some projects, however, have incorporated gender analysis in various aspects of the research process as well. Please note where gender analysis has been used in the work and research process (Bold and mark an X for all that apply. Multiple choices may be selected):

- Research team management and oversight
- Overall research team composition
- Gender-disaggregation in research instrument design
- Overall methodological approach to research
- Conceptual focus of the research issues including differential impacts
- Gendered contextual analysis of stakeholders and key influencers in terms of uptake strategy

- Considered in capacity building efforts of research team

III. Positioning for In-Country Capacity Building

5. Please select the strategies that have been used to ensure built capacity of in-country researchers. (Mark an X beside all that apply. Multiple choices may be selected):

- Mentoring/overlap with principal or lead researchers with in-country researchers
- External advisor or technical support has been brought in that includes mentoring/capacity building
- Research team governance/ management (the extent to which the way the team is set up and roles and responsibilities contribute to capacity building)
- Overall research team composition and structure
- Workshops focused on building key competencies
- Peer mentoring
- Opportunities to expand responsibilities. For example, presentations and co-authorship

6. What would be the strongest indication of success in building in-country research capacity by the end of the program?

IV. Positioning for Uptake and Influence

7. With your understanding of how change and influence work, please specifically name (individual and/or institution) the two most important influencers in your research areas who already exist within your networks

8. With your understanding of how change and influence work, please specifically name the two most important influencers in your research areas who are outside of your existing networks.

9. Your uptake strategy is based on: (*Please number in order of importance only those that apply)*

- Credibility of lead or principal researchers and institutions
- Building on existing relationships and influential networks. For example, key researchers embedded institutionally or in networks
- Fostering new dialogue and relationships to bridge research to influential policy networks/arenas
- Other. Please indicate specifically:

10. Please mark an X beside the three most influential dissemination forums/mediums of your uptake strategy:

- Smaller more pointed meetings with key influencers
- Seminars or conferences specifically for this research
- Piggy-backing existing relevant conferences, events and forums
- Respected journals in the sector
- Our own website

- Key stakeholder websites and/or platforms
- Blogs and Social media (i.e., Twitter, LinkedIn, etc.)
- Press (newspaper, radio, TV)
- Films, CDRoms or Photo exhibitions
- Other (please specify):

11. What would be the strongest indication of research uptake success in your overall research project?

12. Your research uptake strategy is targeted to (please bold the correct response for each group of influencers):

Policy makers a) largely b) in combination with other influencers c) somewhat d) not at all

Practitioners a) largely b) in combination with other influencers c) somewhat d) not at all

Thought leaders/ a) largely b) in combination with other influencers c) somewhat d) not at all

Academics

Please note: Given that research projects are in varying stages of completion, each research project is asked to send copies of their methodology and instrument design by October 31, 2016.

Tool 4: GrOW Key Informant Interview Guide with Principal Investigators

Introduction

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

> The purpose of the interview with principal investigators is to explore the research project strategies in more depth. In particular, the evaluation team will probe some of the issues that have arisen from the project reports, survey results and project officer interviews.

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

<u>These in-depth interview will be tailored to each research project to deepen understanding of specific issues to that project. Some of the following may be used:</u>

1) To what extent has the GROW project allowed you (and your team of researchers) to advance work that may not have otherwise been possible? What unique value-added does this particular funding and project structure provide to your research?

2) How does GROW funding compare to other funding in terms of reporting, learning, iterative support?

3) How has the project navigated the tensions between achieving quality rigorous research, ensuring uptake by key decision makers and building in-country capacity? (possible probing question. There is some debate as to whether uptake should be considered an aspect of quality. From your experience with this research project, what do you think?)

4) What constitutes a quality policy brief?

5) What are the best indicators of success with respect to policy dialogue and engagement?

6) At what stages in the research process should policy makers or other key decision makers be brought in?

7) In your survey (or reports), you indicated that the project will track built capacity of in-country researchers through……… [Probe. Why? Can you elaborate?]

8) What are the most important drivers for building capacity of in-country researchers?

9) What have been your challenges in building capacity of in-country researchers?

10) In your survey, you indicated……………………related to your uptake strategy. [Probe. Why? Can you elaborate?]

11) What are the most important drivers for effective uptake by key influencers?

12) What have been your challenges in designing your uptake strategy and engaging influencers?

13) How do you ensure coherence where different researchers or researchers partners were conducting parts of the research?

14) How do you address the issue of comparability?

Tool 5: GrOW Assessment of Research Project Quality

Introduction

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The purpose of the assessment tool is to review all of the research projects in terms of project quality (rigour/positioning for rigour; legitimacy; importance) regardless of status of completion or methodological risks/weaknesses that may have been identified from the project start. The assessment will attempt to identify the trajectory and progress of each research project as well as trade-offs between elements of quality.

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

		EXPECT TO SEE	LIKE TO SEE	LOVE TO SEE
Quality	Rigour Projects provided: • Methodology, instruments OR • Early research outputs	Research project demonstrates measurement internal validity (minimization of risk and bias errors including incomplete data, sampling, controls, selection bias, endogeneity, randomization errors); external validity (generalizability); reliability (consistency of results across contexts, tools; comparability if appropriate); cogency (authors have considered study limitations and alternative interpretations). A clear link to WEE outcomes. Specific plans for a peer-reviewed article in a named influential journal.	Research project demonstrates measurement internal validity, external validity, reliability, cogency and a clear link to WEE outcomes. Limitations have been rendered explicitly. There is strong coherence between quantitative and qualitative methods (where applicable). Acceptance of a peer-reviewed article in a reputed journal.	There is measurement validity (internal and external), reliability and cogency demonstrated in the methodology or protocol used. Limitations have been rendered explicitly. Clear link to WEE outcomes. There is strong coherence between quantitative and qualitative methods (where applicable). Acceptance to more than one peer-reviewed article in influential journal.

		EXPECT TO SEE	LIKE TO SEE	LOVE TO SEE
	Legitimacy	Ethics review standards (or equivalent) have been approved up to this stage. Indications of gender adaptations and local or contextual adaptation.	Ethics review standards have been approved up to this stage. Demonstration of gender and local adaptations in text. Local researchers played a role in ensuring that study is well-adapted to context.	Ethics review standards have been approved up to this stage. Demonstration of gender and local adaptations in text. Expert advisory group has provided feedback at a strategic time that led to changes in project focus or methodology.
	Importance	A convincing rationale has been made about the relevance, theoretical or methodological innovation and identification of this research gap nationally or globally.	A literature review (not necessarily by the research project) has identified this gap in the literature. Research has been validated by at least one key external stakeholder in the sector.	A comprehensive literature review or systematic review is part of the research project to identify gaps in the literature and ensure importance of this research. Research has been validated by more than one key external stakeholder in this particular sector which may include research team's knowledge of sector.

Tool 6: Assessment of Research Project Strategy Effectiveness

Introduction

The purpose of the assessment tool is to review all of the research projects in terms of project strategy effectiveness or positioning. Since most research projects are in process, this tool will be key to assess the project's management effectiveness, positioning for uptake of key influencers and positioning for built local capacity. The assessment will attempt to identify the trajectory and progress of each research project as well as trade-offs between elements of effectiveness.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

		EXPECT TO SEE	LIKE TO SEE	LOVE TO SEE
Research Project Strategy Effectiveness **(Principal Investigator(s)**	Project Management	Completed quality reporting with adequate specificity of strategies and research progress without unreasonable delays.	Completed quality reporting on time and on budget (or with reasonable delays). Evidence of commitment to quality, research uptake and built local capacity. Strong monitoring and ability to deal with research or team issues.	Completed quality reporting. Evidence of strategic and specific strategies for quality, uptake and built capacity. Evidence of adaptation and learning in project reports. Evidence of ability to mitigate risks and ensure coherence across research outputs.

		EXPECT TO SEE	LIKE TO SEE	LOVE TO SEE
	Positioning for Uptake	Identification of key stakeholders as well as strategies to ensure research is tailored to these stakeholders.	A strategic uptake strategy that outlines specific and key policy, practitioner and academic influencers as well as how the research has been positioned to ensure uptake. Research integrated into an already existing network as well as identification of new stakeholders.	There is demonstration of stakeholder mapping and early engagement. There is differentiation of strategy for different types of influencers. It is clear how different research outputs have been tailored to ensure influence. There is a range of forums or modes of uptake appropriate to different audiences.
	Positioning for Built Capacity	Identification of broad areas, researchers and strategies for capacity development.	Baseline or inventories of skills and capacities. To address these gaps, a variety of strategies and forums have been developed to build capacity. Participation of in-country researchers in key research tasks and responsibilities such as co-authors and presenters.	Baseline or inventory/assessment conducted at start of project to identify gaps. Specific skills or competencies listed. Specificity in terms of how success will be measured in terms of built capacity of in-country researchers. Relatively high participation of in-country researchers as presenters and co-authors.

Tool 7: GrOW Assessment of Research Program Strategy Effectiveness

Introduction

The purpose of the assessment tool is to review the likely impact of the GROW Research Program overall, given its current trajectory, and as more than just the sum of its research projects. This tool will assess the program management, and the likely impact of each of the three program outcomes: enhanced quality of evidence in the sector; enhanced research uptake by key influencers; enhanced built local capacity.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

		EXPECT TO SEE	LIKE TO SEE	LOVE TO SEE
Research Program Likely Impact by End of Program **(Executive Committee**)	Program Management	Timely and substantive feedback to research project partners on project performance and learning. Strong program design for monitoring and learning. Regular and substantive meetings of the Executive Committee to discuss program and partnership level issues.	Strong program design. Timely and substantive feedback. Facilitation of cross and peer as learning. Evidence of adaptation based on timely and strategic monitoring and learning. Evidence of timely supportive process where projects have faced risks internally or externally. Regular and substantive dialogue related to overall program trajectory and impact.	Facilitation of cross and peer as well as program learning. Research project partner largely articulate value-added of grow to their work. A system or process for risk management and non-compliance when a project is not meeting deliverables or experiences a major setback including early trigger mechanisms. Clarity and mutual agreement on success measures/outcomes. Regular and substantive dialogue related to overall program trajectory and impact.

		EXPECT TO SEE	LIKE TO SEE	LOVE TO SEE
	Enhanced body of high quality evidence	Of the research projects reviewed, trajectories indicate a mix of low, moderate and high rigour and legitimacy. Those with low rigour and legitimacy can be improved with corrective measures. Strong and convincing rationales by research projects of the importance of their research for women's economic empowerment.	Of the research projects reviewed, trajectories indicate a mix of quality but the majority are moderate to high in terms of rigour and legitimacy. Indication by the overall program and through evidence mapping (literature and systematic reviews) that GROW research fills identified gaps in women's economic empowerment.	The majority of research projects reviewed are on a path that is highly rigorous and legitimate. Substantiation through evidence mapping (reviews) and key stakeholders) that GROW research fills important gaps in women's economic empowerment.
	Enhanced uptake of research by key influencers	The majority of research projects have identified specific stakeholders and strategies to ensure the research is tailored to them.	For the majority of research projects, there is evidence of tailoring, targeting of key influencers and early engagement. There is also evidence of learning and peer learning related to uptake strategies to become more tailored and effective.	All research projects have demonstrated improvement in their identification of key influencers and strategies for tailoring research products to them. There is early engagement with the majority of projects. The program has helped to provide coherence on success measures and cross-learning.
	Enhanced local research capacity	All research projects have identified broad areas, researchers and strategies for capacity development.	The majority of research projects have done baseline or inventories of skills and capacities. To address these gaps, tailored strategies have been developed to build capacity. Participation of in-country researchers in key research tasks and responsibilities such as co-authoring and presenting at key forums.	All research projects have conducted baseline or inventory conducted at start of project to identify gaps. Specific skills or competencies have been listed. The program has supported peer learning and coherence in terms of how success will be measured. Participation of majority of in-country researchers as presenters and co-authors.

Tool 8: Key Informant Interview Guide with Executive Committee

Introduction

The purpose of the interview guide is to allow Executive Committee members to elaborate specific strategies, areas of success and strength within the program and tensions and trade-offs. The interview results will complement results of the assessment tool to determine the overall likely impact of the GROW Research Program.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

1) What unique contribution does GROW make to the sector of women's economic empowerment? (for each member)
2) What value-added does IDRC bring to the various policy networks and dialogue arenas in women's economic empowerment?
3) What role does the EC play in ensuring quality, uptake and in-country research capacity?
4) How has the program management team (EC) managed consistency and coherence across the diversity of research projects?
5) What are some of the elements that you would consider in terms of assessing the quality of non-experimental research?
6) What are the key drivers (factors, conditions) emerging of research quality?
7) What are the key drivers emerging of effective uptake?
8) What makes a quality policy brief?
9) What are the best indicators of success in terms of policy dialogue and engagement?
10) What are some of the key drivers emerging of effectively building local capacity?
11) What might be considerations in the assessment for the different governance structures: north-south partnerships; south country based; south-south partnership
12) What corrective measures or challenges in the overall Research Program do you feel could be mitigated by policy or monitoring changes at the Program (EC) level?

Questions specifically for the IDRC Program Manager:

1) What is the structure, frequency and process for program level monitoring, learning and reflection?
2) Where are the example of success/strength/positive deviance with respect to effective uptake strategies?
3) Where are the example of success/strength/positive deviance with respect to building local capacity?
4) What are the processes for early detection and risks in the program with respect to non-compliance or other risks?
5) Can you comment on research projects that were originally approved but were discontinued?
6) Any additional hindsight reflections or comments that would be helpful for the evaluation team?

Thank you for your time and consideration

Tool 9: Sectoral Evidence Mapping Framework

Introduction

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

The purpose of the mapping framework is to illustrate, schematically, where the GROW research is situated in research on women's economic empowerment, particularly related to outcomes. The framework provides a structure to review key sites, systematic reviews and sources of research on women's economic empowerment. The result of the evidence mapping will be a sectoral schematic mapping of current WEE research including GROW research. Note: The scope of this evaluation does not permit an exhaustive systematic review. Rather, the review will draw on existing systematic and literature reviews on the sector as well as key global platforms.

WEE INTERVENTIONS	OUTCOMES							
	CHALLENGES OF WEE							

Note: Systematic reviews both within GROW and outside of GROW will be important here.

Tool 10: Key Sectoral Stakeholder Informant Interview Guide

Introduction

The purpose of the mapping framework is to show where the GROW research is situated and adds value to existing women's economic empowerment interventions and studies, particularly related to outcomes. The framework provides a structure to review key sites, systematic reviews and sources of research on women's economic empowerment.

The Growth and Economic Opportunities for Women program (GrOW) is a five-year, multi-funder partnership with the UK's Department for International Development (DFID), The William and Flora Hewlett Foundation, and the International Development Research Centre (IDRC).

The Universalia Management Group has been mandated to conduct a mid-term evaluation of GrOW. The evaluation will examine the current portfolio of research projects funded through the program to examine research quality in terms of positioning and importance, research effectiveness (positioning for uptake and built local capacity) and likely impact of the overall research program. The evaluation is formative and aims to identifying lessons learned and areas for improvement for program-wide benefit going forward. Any specific questions about this process can be directed to Nanci Lee, Team Leader: nancilee@eastlink.ca

[In advance of the interview, provide the summary document outlining the research output summaries]

1) Describe your position and an overview of your experience in the sector of women's economic empowerment
2) What is your perspective on the value-added of GROW research to the sector overall?
3) Is there anything in particular, about these funding bodies and this structure that might add value to the research, policy or practice?
4) Is there a key gap in the sector of women's economic empowerment that you would have liked or expected could be addressed by this Research Program?
5) [Probe] What are the particular elements that you feel add value? Are they in terms of topic, methodology, key influencers.
6) Can you comment on legitimacy of the research agencies [identified relevant to stakeholder geographically or topically] involved and their reputation in the sector?
7) Do you have any advice or feedback for enhancing the overall research program since it is in process?
8) Do you have anything else to add that may be helpful for the evaluation team to know?

Thank you for your time and consideration.